Classification of Galvanneal Steel using Optical Texture Analysis

Scott Lee Woodham

A thesis submitted in partial fulfillment of the requirements of the University of Glamorgan for the degree of Doctor of Philosophy

August 2000

University of Glamorgan

ABSTRACT

Off-line Galvanneal steel classification using chemical and mechanical processes, has matured to a level where they are now able to confidently assess the state of iron/zinc alloying within the galvanneal coating. However, the assessments require that the samples be extracted and analysed off-line. This restrictive requirement means that the techniques tend to be destructive in nature, leading to unnecessary wastage. Moreover, they are unable to provide immediate line control feedback to ensure that the alloying condition is consistently at the optimum level.

Application of image processing methods to the analysis of the steel surface morphology presents the ability to classify coating quality based on extracted texture information. Using images extracted from actual galvanneal steel coils, the primary objective is to investigate the link between the annealing phase of the steel surface and the digital optical image equivalent. This investigation offers the potential of developing and implementing an on-line real-time annealing phase classification system that is less intrusive and time consuming than the existing laboratory analysis techniques.

Known image processing procedures have been applied to off-line galvanneal steel image samples to determine the degree of correlation between the captured image texture and the level of surface alloying. The procedures include wavelet compression, first and second order statistical pattern recognition techniques, genetic algorithm and K-nearest neighbour classifiers (Knn) and the rank-conditioned (R-C) morphological transform. Two novel techniques have emerged from the investigative research, a grey level co-occurrence image filtration procedure and a morphological template optimisation scheme. The combination of the R-C transform using spiral templates with the image filtration system produced classification rates of 100% and 80% for two sets of galvanneal data, despite the latter being affected by high levels of noise. The second order statistical approach with the Knn classifier produced results of 93% and 62% respectively.

Further galvanneal test images should increase the confidence in the applied morphological techniques prior to their application as an on-line classification system.

DECLARATION

I declare that this thesis has not been, nor is currently being, submitted for the award of any other degree or similar qualification.

Signed: 

Scott Lee Woodham

ACKNOWLEDGEMENTS

I wish to thank my Directors of Studies: Prof. Robert Williams for initiating the study and offering me a chance to fulfil a long time ambition, and special thanks must go to Dr. John Rees for sharing his vast knowledge of image processing techniques, and for defining the core structure of the work undertaken.

This project would not have been possible without the research staff at British Steel, Port Talbot. I would like to express my gratitude to one and all, especially Vivian Worner for supplying the image samples and for detailing the processes involved in the galvannealing operation.

I would also like to thank Alison Evans, Russell Davies, Leighton Spurgeon, Andrew Beer, Chris Stanford and Simon Ho for their continued friendship and patience.

Above all, I am eternally grateful to my Parents Barbara and Tony Woodham for their moral and financial support throughout my studies, and to my Grandfather Cecil Powell for his constant encouragement and for his remarkable, waist-expanding culinary creations.

List of Contents

List of Figures

List of Tables

SUMMARY OF APPLIED IMAGE PROCESSING TECHNIQUES

(i) Wavelet transform [1] - used to compress data for the removal of elements between the galvanneal steel crystals. It was unable to isolate any class specific features and offered no improvement over the quality of the standard images. Led to the following technique of extracting and exploiting the differences in pixel intensity between each annealing class.

(ii) First order statistical technique [2] - analyse image intensity. The histogram procedure was too sensitive to the value of an individual pixel and produced results that failed to indicate any form of in-class image homogeneity. This influenced the choice for the next method of examining the crystal structures through the interaction of pixel-pairs.

(iii) Second order statistical technique [3] - evaluate structural information. The grey level co-occurrence matrix (GLCM) is less affected by the value of an individual pixel. It attempts to model the surface texture based on the variations in pixel-pair intensity over the entire image. At 100%, the technique produced a promising set of a priori results but failed to produce tight, isolated clusters to support classification without any prior knowledge. The results generated by the co-occurrence matrix were analysed and categorised using the following two classifiers.

(iv) Genetic algorithms [4] - examine how the second order results perform in a non-a priori environment. Genetic algorithms (GA) were used to reduce cluster variance and improve group separation by implementing an element-weighting scheme. This method was unable to rectify the problem of cross-cluster contamination.

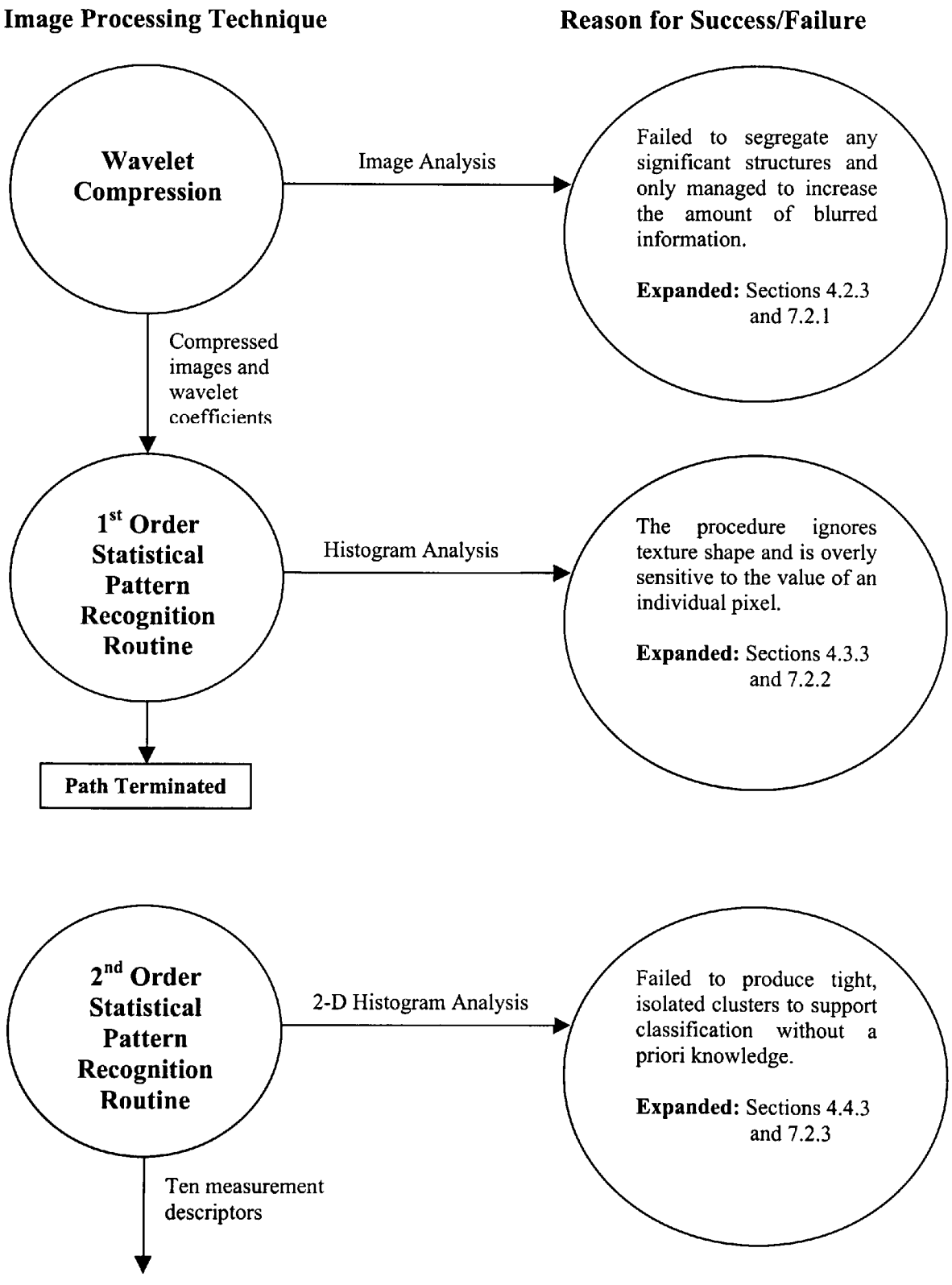
(v) K-nearest neighbour [5] - evaluation of the second order results using an alternative classification routine. The K-nearest neighbour (Knn) classifier defines the class of an incoming data point based on its K nearest neighbours. This routine was incapable of producing tight clusters and was affected by the order in which the test data was analysed.

(vi) Rank-conditioned morphological transform [6] - using templates of varying size, shape and height, the crystal formations on the surface of the steel substrate was examined for any class-specific boundary shapes. At best, it managed to classify 71% of the galvanneal images but was unable to extract information from or around the blurred regions. Pre-processing techniques to improve probe access and increase the rate and accuracy of measurement extractions, were therefore developed.

(vii) Novel grey level co-occurrence filtration scheme - "opened up" blurred regions to enable an increased number of measurements whilst preserving basic crystal shape. In conjunction with spiral templates at varying height, the scheme improved the rate of classification to the overall maximum of 80%. However, the optimum height of the analysis template was difficult to determine. Consequently, methods for the determination of the height of a probe were investigated.

(viii) Novel approach to the design of template prototype intensity using the grey level co–occurrence matrix - the grey level co-occurrence matrix was used to analyse the pixel data encountered by each analysis structure using sample training images from each galvanneal class, with the height of the template set according to the significance of the texture information extracted. The optimisation routine generated at best a classification result of 66%, an improvement over the 62% of the GLCM and Knn combination, but offering lower classification rates than the template intuitively designed by an expert user.

Progress Path and Interrelationship of the Applied Techniques



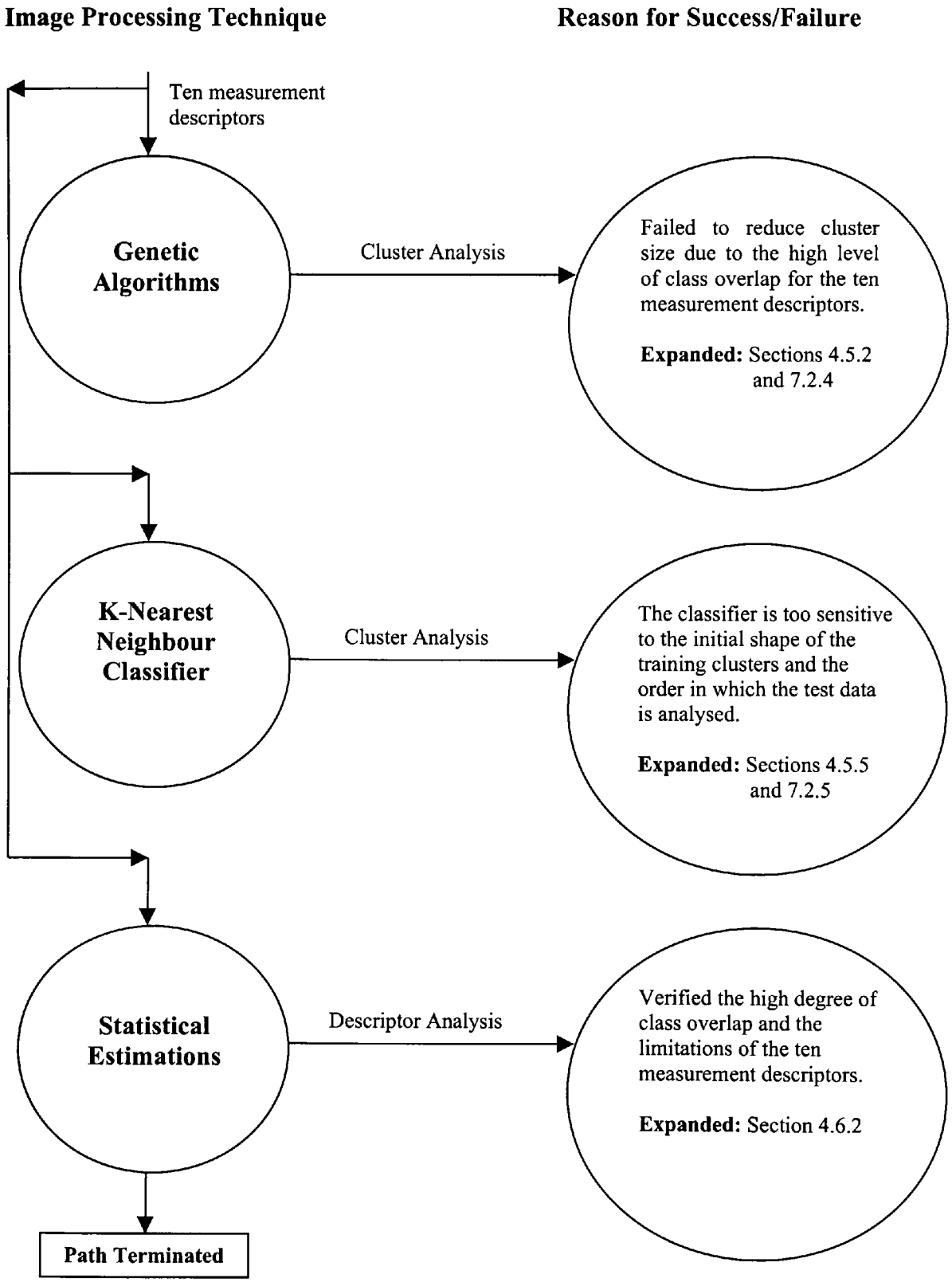


Image Processing Technique
Reason for Success/Failure
Ten measurement descriptors
Genetic Algorithms
Cluster Analysis
Failed to reduce cluster size due to the high level of class overlap for the ten measurement descriptors.
Expanded: Sections 4.5.2 and 7.2.4
K-Nearest Neighbour Classifier
Cluster Analysis
The classifier is too sensitive to the initial shape of the training clusters and the order in which the test data is analysed.
Expanded: Sections 4.5.5 and 7.2.5
Statistical Estimations
Descriptor Analysis
Verified the high degree of class overlap and the limitations of the ten measurement descriptors.
Expanded: Section 4.6.2
Path Terminated

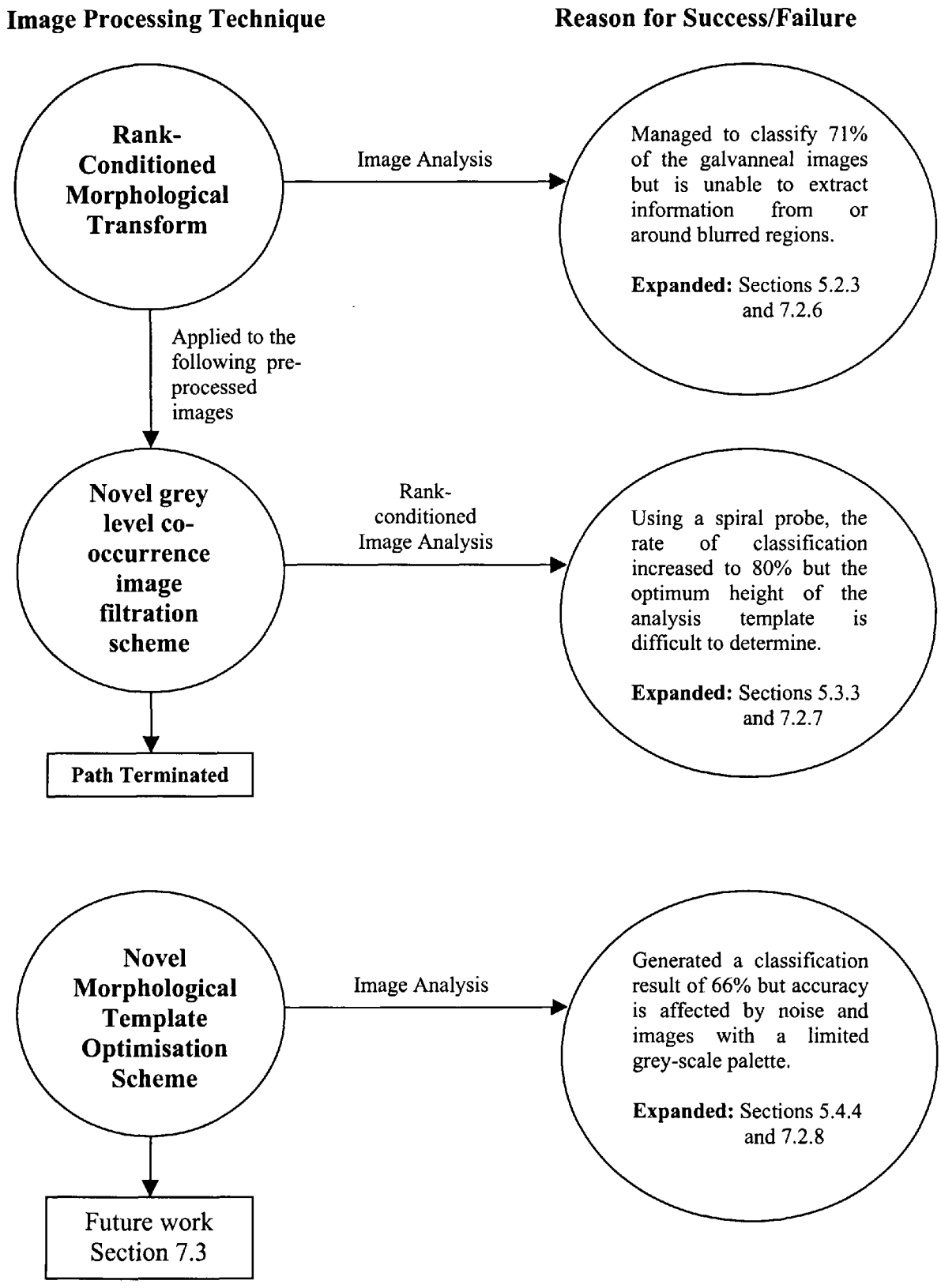


Image Processing Technique
Reason for Success/Failure
Rank-Conditioned Morphological Transform
Image Analysis
Managed to classify 71% of the galvanneal images but is unable to extract information from or around blurred regions.
Expanded: Sections 5.2.3 and 7.2.6
Applied to the following pre-processed images
Novel grey level co-occurrence image filtration scheme
Rank-conditioned Image Analysis
Using a spiral probe, the rate of classification increased to 80% but the optimum height of the analysis template is difficult to determine.
Expanded: Sections 5.3.3 and 7.2.7
Path Terminated
Novel Morphological Template Optimisation Scheme
Image Analysis
Generated a classification result of 66% but accuracy is affected by noise and images with a limited grey-scale palette.
Expanded: Sections 5.4.4 and 7.2.8
Future work Section 7.3

Chapter 1

Introduction

1.1 Introduction

There has been an increasing interest in hot-dipped iron/zinc (Fe/Zn) alloy coated steels (Galvanneal) in recent years, particularly from high volume markets such as the motor industry [7]. Changes in market pressure have placed an increased demand on companies within the steel manufacturing sector to raise the standard of the product in terms of surface quality, resistance to powdering, corrosion, and weldability [8, 9, 10]. These requirements dictate that imperfections in the galvanneal strip must be recognised at the time of manufacture and the accompanying process variables must be noted in an attempt to reduce the probability of similar imperfections appearing in any succeeding strips. At present this is a manual operation and is prone to human error.

The quality of galvanneal with regards to physical and coating properties is dependent upon a number of process variables. These include zinc bath chemistry, surface properties, coating conditions, furnace annealing and the strip cooling temperature cycles [7, 11]. The successful control of these parameters, assessed by the actual galvanneal

output condition, holds the key to producing consistent annealing levels in the alloy coated steel strips.

1.2 Galvannealing Process

Steel is subject to corrosion under normal conditions of use. In order for a steel strip product to offer a lifetime of satisfactory service, a sacrificial barrier needs to be placed between the steel and the harmful elements of the outside world. One such barrier commonly used to protect the product consists of an iron/zinc alloy coating formed using a method known as galvannealing.

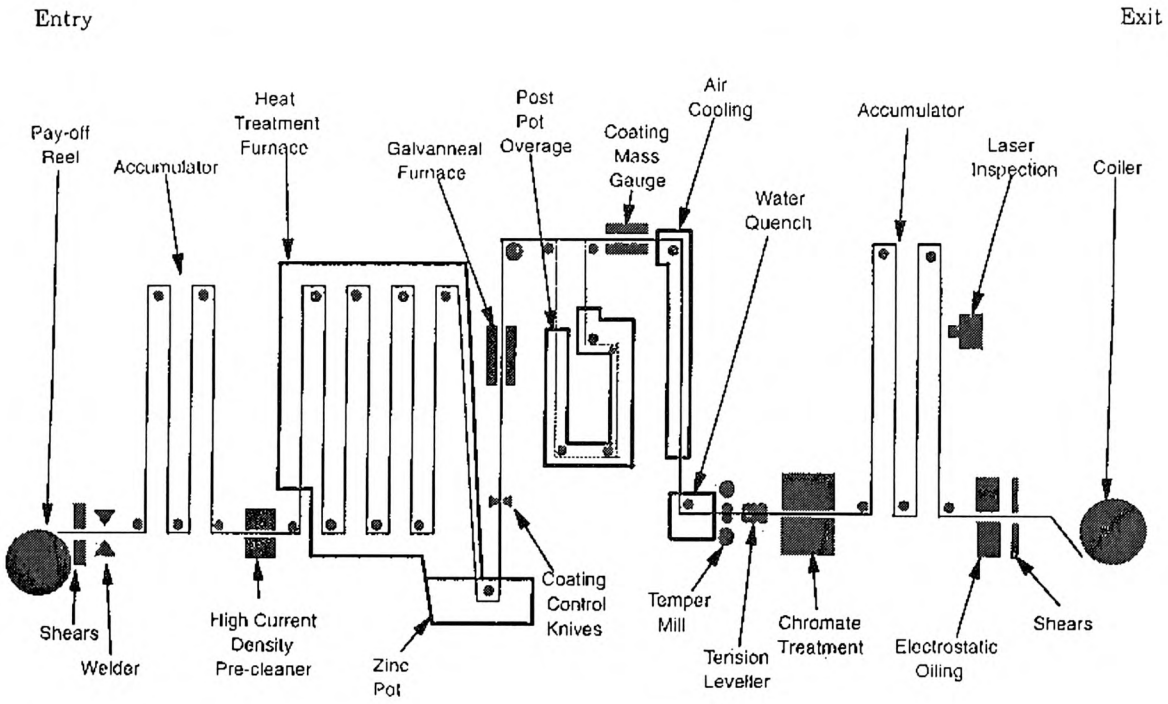
The galvannealing process provides a substantial increase in corrosion resistance when compared to the standard bare steel surface [11]. Steps involved in creating a galvanneal strip of steel initially follow the same route as the galvanising operation but with the inclusion of what is recognised as an annealing phase.

Galvanising [12] is achieved using one of two methods. The first method, electro-galvanising involves the use of an electro-chemical reaction to fuse the alloy mixture to form a strong bond with the steel surface. The second method, used to form the collection of examined test samples, is known as hot dip galvanising, abbreviated to HDG [13].

The accumulator behaves as a storage station enabling the rest of the process to continue whilst the cutting and welding of the two strips is performed (See Figure 1.1). The process begins with the cold-rolled strip being welded to the tail edge of the previous coil and directed towards the accumulator. Next, the steel is passed into the heat treatment furnace and is heated to approximately 680° C to remove any oil residue and unwanted

surface material. It then enters the annealing section of the furnace and is further heated until it reaches a temperature of around 730° C. This causes the strip to re-crystallise before it is rapidly cooled and sent to the zinc bath which is maintained at a constant 465° C.

Figure 1.1: Modern Hot-Dip Galvannealing Line [11]



Small amounts of aluminium are added to the molten bath of metal to restrict the thickness of the iron/zinc coating. The quantity of the zinc on the surface of the steel is further controlled using a series of air or steam jets, known as the coating control knives.

The annealing stage involves passing the galvanised strip through an induction heating system immediately after the coating operation. The strip is then annealed at about 500° C for approximately 15 seconds, to form a long lasting, protective layer and is prepared for the final stage of manual inspection [11, 14].

1.3 Objectives

Using images extracted from actual galvanneal steel coils, the primary objective is to investigate the link between the annealing phase of the steel surface and the digital optical image equivalent, for possible use as an automated on-line galvanneal steel classification system.

The principal goals of the research can therefore be defined as follows:

(i) The implementation of suitable feature extraction algorithms and corresponding classification methods.

(ii) The application of the algorithms to various steel samples to assess the degree of correlation between the steel image texture and the surface alloy properties.

(iii) Evaluation of classifier capability as a method of steel coating analysis.

(iv) The refinement and optimisation of the applied feature extraction and analysis routines to enhance the rate of classification.

The secondary objectives of the project are required for the comparative evaluation of the developed techniques:

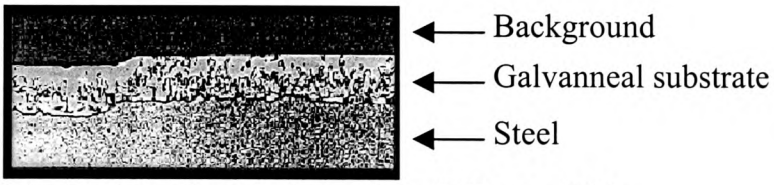
(i) The assessment of the ability of the investigated techniques to classify alternative forms of natural texture other than the galvanneal steel crystals.

(ii) Formal comparison with proven image processing techniques.

1.4 Galvanneal Coating Condition and Image Quality

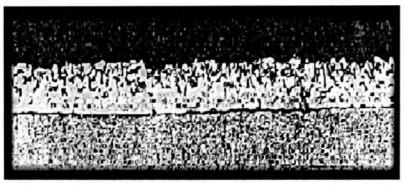
Off-line steel samples were captured using the Leco 2001 image analysis system. This is comprised of a high magnification optical microscope (maximum resolution x500) with both black and white film and CCD camera attachments. The images were acquired at three different levels of magnification (x50, x200 and x500) and divided into classes of annealing based on the level of iron (Fe) present in the steel coating [8, 14, 15]. For basic descriptive purposes, each class can be either defined as under-alloyed (Fe<=9%), over-alloyed (Fe>=11%) or alloyed at the optimum state (Fe>9% and Fe<11%). Further sub-divisions can be performed based on the Fe percentage level but their annealing type is restricted to one of three labels. Figure 1.2 presents a cross-view (x1000) of the galvannealing process at various stages of annealing and illustrates the three distinct groups.

Figure 1.2: Cross-section of the main classes of galvanneal steel, Magnification level x1000 [11]



(a) Under-alloyed strip of steel (Fe <= 9%)

(b) Over-alloyed strip of steel (Fe >=11%)



(c) Cross-view of a strip of steel annealed to the optimum state (Fe >9% and <11%)

The first example, Figure 1.2a, demonstrates the case of an under-alloyed strip of steel. The iron/zinc coating has not formed the desired stable layer of crystals and the process has softened the outer bed, giving rise to a smooth surface with a high reflectance value. The coating has become what is known as very ductile and the protective substrate has affected the underlying steel surface.

The over-alloyed strip in Figure 1.2b presents a more or less undamaged steel surface at the expense of the iron/zinc coating. The coating has become very brittle and has caused large fissures to appear in the protective layer. These breaks undermine the sacrificial corrosion control properties of zinc, which normally provides the extra protection by corroding in preference to the steel.

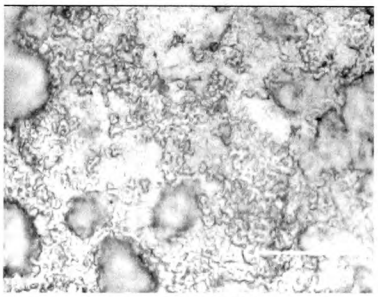
Finally, the optimum state in Figure 1.2c displays definite iron/zinc crystal formations and has established a strong bond with the underlying steel. No fractures are visible on the galvanneal coating and the steel base remains intact.

For an image processing based galvanneal inspection system to operate as a potential on-line classifier, the digital samples should be captured from a realistic viewpoint that offers an accurate representation of the qualities that define each annealing type. Extracting side-view images from the steel product as it traverses the galvannealing line is not

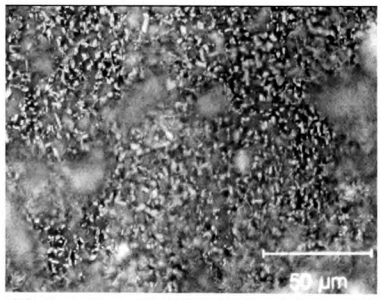
practical. The level of movement in the vertical plane coupled with the etching and polishing required for the strip to be prepared for high level magnification capture and analysis (from this angle), complicates the classification task. Similarly the side-view would not reveal the annealing variations across the coating surface. Therefore, all the test samples have been captured using a top-down perspective.

The following surface-view images (Figure 1.3) are samples drawn from the x500 galvanneal collection and are sub-divided into five known calibrated coating conditions.

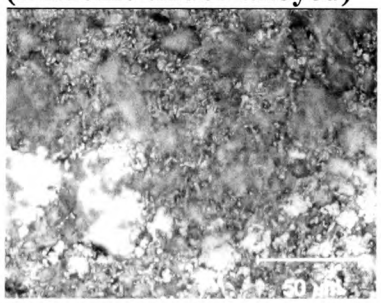
Figure 1.3: Examples taken from the five classes of galvanneal steel (surface-view), 499x386, 8-bit pixel depth, Magnification level x500



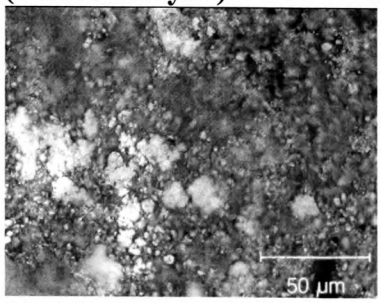
37.6 g/m, 5.38% Fe, 0.66% Al (Extreme under-alloyed)



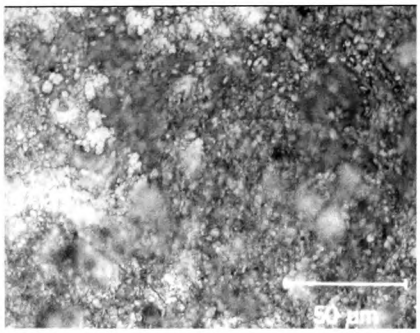
40.3 g/m, 8.42% Fe, 0.61% Al (Under-alloyed)



41.9g/m, 10.79% Fe, 0.59% Al (Optimum-alloyed condition)



42.7g/m, 12.08% Fe, 0.6% Al (Over-alloyed)



48.2g/m, 14.37% Fe, 0.54% Al
(Extreme over-alloyed)

From an image capturing perspective, the surface viewpoint offers a more practical solution than the side-view. For the purpose of classification however, it still poses several problems. There is no clear indication of any surface defects in the form of micro-cracks and the view is unable to reveal any effect on the underlying steel surface. The classification challenge is further enhanced through a problem of depth of field blurring caused by limitations with the optics used in the Leco 2001 image capture system. As the optical lens approaches its magnification limit, the system loses its ability to represent clearly the crystal formations. This results in a blurring of up to 97% of the image in certain cases, and has the effect of restricting the ability of an image feature representative to accurately define a class. The problem is one of generating a classification feature that is indicative of the surface properties within each annealing group and capable of confidently distinguishing between each alloyed state.

Early image analysis using wavelet compression, first and second order statistical analysis with genetic algorithm and k-nearest neighbour classifiers (Chapter 4), failed to yield a useful measure of a sample's coating condition from its digital representation. Similarly, basic evaluation techniques such as intensity assessment and the analysis of pixel-pair interactions, served only to highlight the obvious differences between the extreme under-

alloyed class (5.38% Fe) and the four remaining groups. However, the combination of the rank-conditioned morphological operator with the grey-level co-occurrence filtration scheme (Chapters 5 and 6), which offers the ability to analyse both crystal size and 3-D texture profile, has generated classification rates that may allow adequate discrimination for practical use.

1.5 Project Justification

It is essential that the outgoing steel morphology is always at the optimum annealing state. Currently if a defect is found in a coated strip, then the entire section is wasted. An automated classification system would ensure that the quality of the galvanneal product is consistently at the desired standard required by the customer. The conversion from a manual to an automatic system would also introduce a number of other, more obvious benefits, such as less wastage, cheaper operational costs and an increase in product assessment reliability.

1.6 Original Contribution to Knowledge

The work undertaken in the project represents a systematic approach to the investigation of the practicality of the automation of galvanneal output quality inspection using surface digital imaging.

The original contribution to knowledge is based on the application of established and novel techniques, to an area of work in which these methods have received relatively little attention. Additionally, Chapter 5 introduces two new areas of research involving both the

grey-level co-occurrence matrix and a variation on the rank-conditioned morphological transform. More specifically:

(i) Section 5.3 defines an image filtration scheme based on the grey-level co-occurrence matrix (GLCM). The matrix offers an indication of the degree of depth of field blur affecting the galvanneal steel crystals. By filtering these blurred areas within the GLCM and transferring the results back to the original source image, crystal shape and structure is retained whilst the noisy regions are more accurately examined using the rank-conditioned morphological process (Chapter 5). The process is defined mathematically in section 5.3.2 and applied to the x500 samples in section 5.3.3. In segment 6.2.2 the technique is applied to the x200 image set in addition to other examples of natural texture.

(ii) Section 5.4 defines a morphological template optimisation scheme based on the grey-level co-occurrence matrix. A number of GLCMs are generated using each probe element as individual displacement distance and angle parameters (Figure 5.9). From this data, it is possible to examine the grey-tones encountered by each element as applied to an appropriate training image set, and to therefore determine an optimum intensity threshold value for the complete image collection. Thus maximising useful data containment for any shape or size of probe on a class by class basis. The process is defined mathematically in section 5.4.3 and applied to the x500 and x200 sample sets in sections 5.4.4 and 6.2.2 respectively.

1.7 Interim Conclusion

The surface condition of the steel coating changes as the annealing operation is carried out. Variations in the process variables, causes the formation of distinct iron/zinc phases which are not detected until the annealing phase is complete.

This creates a situation where not only is the task of quality control difficult to perform, but the detection of the optimum coating variables for controlling the annealing environment is also difficult to assess.

Therefore, the main objective is to classify the anneal state of coated steel strip products, using surface images classified according to their optical texture. A successful project would offer the underlying technology necessary for the development of real-time feedback, leading to the possible generation of automatic line-control data.

To ensure the validity of the results, the novel methods are applied to standard images for evaluation against established techniques.

1.8 Chapter Outline

Chapter 1 "Introduction" - introduces the area of steel strip quality control, offers an overview of the annealing process and defines the objectives.

Chapter 2 "Literature Survey" - reviews the current publications within the field of study and determines the position of this particular project.

Chapter 3 "Test Data Preparation" - introduces the galvanneal image set and the methods employed to prepare the source data prior to analysis.

Chapter 4 "Galvanneal Steel Image Analysis using Established Image Processing Techniques" - examines the data set using established methods of image analysis. It also raises the question of whether the images are of such a poor quality that their use means that successful galvanneal steel classification is an unattainable goal, or whether the applied procedures are unable to represent the steel crystals in a suitable manner.

Chapter 5 "Galvanneal Steel Image Analysis using the Rank-conditioned Morphological Transform and Novel Image Processing Techniques" - sees the application of the rank-conditioned morphological transform, together with a novel grey-level co-occurrence image filtration and template optimisation scheme and the highest overall rate of classification at 80%.

Chapter 6 "Final Analysis and Discussion" - analyses the Brodatz and x200 magnified galvanneal image sets using the techniques outlined in Chapters 4 and 5. The results indicate with some assurance that image quality is the reason for the poor classification results rather than the techniques used. Although further galvanneal samples at x200 and x500, will increase the confidence in the conclusions drawn.

Chapter 7 "Summary and Suggestions for Further Work" – summarises the limitations of the techniques investigated for galvanneal classification and defines a definite progress path that offers both short and long-term options.

Chapter 8 "Conclusion" – summarises the findings and discusses the viability of an on-line automated galvanneal steel classification and line-control system.

Chapter 2

Literature Survey

2.1 Introduction

Little research has been conducted into the specific area of galvanneal coating assessment through image processing techniques. Actual publications that offer evidence of image processing and galvanneal coating research remain scarce. The major body of work in the field relies primarily on mechanical and engineering methods, typically depend upon a certain degree of human intervention.

2.2 Existing Methods for Galvanneal Classification

The current methods employed by the steel industry revolve around a combination of galvanneal environment sampling and human visual inspection, coupled with off-line laboratory analysis.

The computer system offers real-time sampling of the annealing environment and not only provides a continuous report of the important process variables but also offers the ability to control the power, line speed and temperatures according to the feedback provided by the on-site laboratory steel coil assessments. Moreover, the apparatus

maintains a vigilant watch over the running of the entire procedure, and detects and alerts the operators to the location of any operating faults.

As the galvanneal strip leaves the exit accumulator (Chapter 1, section 1.2, Figure 1.1), it passes through a manned inspection point. From a classification perspective, the room consists of a number of monitors that provide the operator with a variety of alternative viewpoints of the passing steel product, namely the upper and lower surfaces. To allow for a closer, more detailed examination of any noticeable defects, the strip can be paused and inspected prior to being cut and coiled.

Once the operation is complete, samples are extracted from the processed coils and the coating quality is assessed on-site using a number of techniques. These techniques include [7] surface roughness determination, X-ray diffraction/flourescence, Mössbauer spectroscopy [16], coulometric analysis, assessment of adhesion – powdering and flaking, and measurement of surface reflectivity.

Koesveld et al [17] use the roughness test to assess surface reactivity by analysing the frequency and size of crater formations. They also combine X-ray diffraction with a process known as anodic dissolution (coulometric analysis), whereby the galvanneal coating is progressively stripped away until its corrosion potential reaches a characteristic plateau, the result is a graph that defines the thickness of each annealing phase. Koesveld et al conclude by stating that the combination can be used to identify the various Fe/Zn constituent phases within the galvanneal layer. Grant and Cook [18, 19], compare the accuracy of galvanneal phase identification when using X-ray diffraction (XRD), X and gamma ray Mössbauer spectra (XMS and GMS respectively), and conversion electron

Mössbauer spectroscopy (CEMS). They state that the GMS measurement was very good at determining the phases formed in the galvanneal coating by probing through the entire protective layer to the underlying steel surface. Lee and Hiam [20] use coulometric analysis to assess the corrosive resistance of galvanneal steel. Mataigne et al [21] also use coulometric analysis, producing stripping curves that are representative of the under, optimum and over alloyed classes. They also perform surface friction analysis through flaking resistance (assessment of adhesion) to examine the quality of the steel substrate. In [8], Gray et al assess the effect of coating weight and iron content on the level of surface adhesion using the one-half inch compression test. Arimura et al [22] examine the hardness of the galvanneal layer using a three point and a 60° bending test. Leroy [23] applies the V bend and cup test to various steel alloy products, and illustrates the correlation between powdering loss and the level of iron present in the protective coating. Hashiguchi et al [24] deform the original galvanneal sample using either stretch forming, plane strain forming or deep drawing to compare the various effects of steel sheet deformation. After the deforming process is complete, the specimen is re-flattened and adhesive tape is applied to the surface. Once removed, the tape is analysed using X-ray flourescence for any variations in the amount of powder/flakes removed. In addition to finding differences between the methods of deformation, Hashiguchi et al confirm the results presented in [8, 22, 23] and state that the coating weight and Fe content effects the level of surface adhesion and therefore the quality of the galvanneal coating. Watts [7] applies a number techniques to assess the condition of galvanneal specimens during and after annealing. One particular technique involves the extraction and examination of surface reflectivity measurements. The light from a halogen lamp is passed through a focussing unit and directed onto the sample at 45° where upon, the amount of reflected electromagnetic radiation is measured using a light sensitive silicon photodiode. Based on

the results produced, Watts states that the apparatus has the potential to be used for on-line monitoring and control of the level of iron /zinc alloying. In [25], Badger and Pyle, describe the operating performance of the "Isys is-2000" automatic surface inspection system. Using four CCD cameras the system captures various views of the upper and lower steel surface. By performing image processing functions and library comparisons, it is able to detect surface faults to an accuracy of 100% and is able to classify those defects at a rate of 60%.

To date, the majority of processes applied to the galvanneal steel surface to ascertain the quality of the sacrificial layer, remain destructive measurements that are unable to provide immediate feedback for line control adjustment. Nevertheless, they are well researched and tested and provide reliable results given enough time.

2.3 Image Processing Nomenclature

The rapid development of computer technology in terms of raw processing power and ease of use, coupled with the falling costs of ownership, has led to an explosive increase in the number of applications developed within the field of image processing and analysis [6, 26, 27]. The purpose may vary between each application but the overall goal remains the same - to mimic certain aspects of human vision and perception, and where possible or necessary improve upon the process, e.g. for night vision, hazardous or polluted environments, or for under-water scene analysis.

Van der Heijden [26] verifies the common goal by outlining five points that define the framework of an image based measurement system. These points are equally applicable to both the human and computer approaches to image processing and analysis:

(i) Object detection

Scan the scene for any objects of interest.

(ii) Object classification

What label can be assigned to the object under consideration?

(iii) Parameter estimation

Define the object's size, position and orientation.

(iv) Shape analysis

Determine an object's shape and geometry.

(v) Scene description

Examine the objects in the scene and determine their relationship.

Van der Heijden [26] further states that the main application of image based measurement systems for industrial environments e.g. a typical galvannealing line, is that of quality control and assessment.

Haralick and Shapiro [27] define the term image processing as, the various operations that can be applied to digital image data. These operations include, image compression, restoration and enhancement, spatial filtering and computer vision techniques [26]. They then proceed to characterise an image as the spatial representation of an object, a two-dimensional or three-dimensional scene or another image. The description is then refined and expanded in the context of computer vision and understanding. An image is defined

as a matrix of numeric digits, each representing a quantized intensity value, based on the level of reflected light detected on a source object using the digital image formation system [28].

Pitas [28] details the steps involved in the conversion between the desired scene/object and the digital image equivalent. The acquisition system comprises of an optical lens, a sensor and a digitiser (or frame grabber). He states that the apparatus used to convert an optical signal into a digital signal introduces the problems of distortion, deformation and degradation. Haralick and Shapiro [27] repeat the same inherent problems observed by Pitas in more fundamental terms by merely commenting on the difficulty in producing a computer vision system when the units of observation do not match the units of analysis. They then refer to the individual units of an observed digital image matrix as being a pixel (pel [29]). Each pixel or picture element [29] can be defined using two properties [27], the first is the row-column position within the matrix, and the second is the grey tone (grey scale/level) value that defines element intensity. These two parameters help to define which image, object, or scene is perceived within the captured boundaries. Moreover, the specific parameter settings, coupled with the mode of pixel interaction and organisation, loosely define what humans would perceive as texture.

A universally accepted term for texture has yet to be formulated [28, 30, 31, 32, 33]. Nevertheless, the descriptions that are currently in existence contain common phrases and elements that define the term to an acceptable level. Modestino et al [34] claim that at the time of publication (1981), the most widely accepted definition of texture was written by Picket [35], Hawkins [36] and Pratt [37]. They collectively state that texture is a collection of homogeneous or quasi-homogeneous pixel patterns that are repeated

periodically or semi-periodically throughout a perceived image region. Haralick [38], and more recently Dentith [39], adds a further dimension by describing texture as the spatial organisation, relationship and distribution of grey-tones. Pitas [28] and Gonzalez and Woods [29] adopt the modern approach to defining texture, by incorporating specific phrases that conjure up ideas of how the material object must physically feel as well as its appearance as a digital image, employing property descriptors such as texture smoothness and coarseness. Care must always be taken when defining texture in such a way, since an object that appears smooth to the naked eye may exhibit coarse characteristics when viewed at some magnification. Haralick and Shapiro [27] provide an updated description that encompasses all the previous definitions of periodicity and homogeneity and supplements them with modern terms such as, density, roughness, and directionality to produce a more complete and condensed representation of texture in the context of computer vision. They then proceed to define the unit of texture as being the texture element, abbreviated to texel. A texel (unit pattern [40]) is perceived as possessing three variable parameters, a row-column location index, the size of the texture element and finally the vector of texture properties formed from local, homogeneous pixels. The term has been utilised throughout Chapters 4 and 5 to represent isolated steel substrate crystal formations that are repeated either periodically or semi-periodically. Collectively, the texels (unit patterns) form class-specific pattern characteristics that may be isolated and measured using various feature extraction and analysis techniques, such as those outlined in the fourth and fifth chapters. A feature [27] is defined as a measurement that is designed to contain as much information as possible about image/object grey-tone intensity, texture, structure and shape, for the purpose of image/object classification. Extracting and comparing feature measurements reduce the number of comparisons required for the classifier to assess and assign the image/object to the determined

category. The process of selecting the appropriate feature representative and feature classifier therefore, remains an important issue with any computer vision and analysis system. Failing to apply the correct extraction and evaluation schemes, weakens the entire classification operation and limits the accuracy of the final results.

2.4 Steel Inspection and Texture Analysis using the Investigated Techniques

The image processing approach may not offer the same direct analysis of the galvanneal surface as existing industrial methods, but the procedures involved tend to be less destructive and less intrusive.

In [41, 42], Wiltschi et al attempt to produce an automatic assessment scheme for the inspection of non-coated steel using image processing techniques. Current methods of manual quality control follow the same path as galvanneal steel analysis, i.e. off-line samples extracted from main coil (are etched and polished), analysed under a microscope and finally, the surface quality is manually compared against standard samples illustrated on a fixed chart. Using mathematical morphology, Gabor filtering and spatial and frequency feature extraction procedures, they offer a system that attempts to mimic the operation of a trained metallurgist. The system was designed to isolate, extract and classify characteristic carbide structures. Attaining a classification rating of between 60% and 70% for the test image set (depending upon the classifier and distance measurement utilised), they claim that the image processing approach offers a level of performance that is comparable to that of the trained human examiner.

The use of wavelet compression [43, 44, 45, 46] for galvanneal inspection was an exercise in steel crystal isolation, rather than surface texture analysis. The process (Chapter 4, section 4.2.3) attempted to weaken the effect of noise introduced by the depth of field blur, whilst retaining basic crystal boundary shape and structure.

Öktem and Egiazarian [47] continue their work first introduced in [48] and [49] by applying the wavelet compression routine to three test images that had been corrupted by "film-grain type noise". They state that noise affects the performance of any compression technique by firstly, interfering with the level of neighbouring pixel interaction and secondly, by introducing undesirable pixels that must still be encoded even though they are not part of the original scene. Using the higher resolution wavelet coefficients and filters, coupled with a basic quantisation scheme, Öktem and Egiazarian, endeavour to target and remove the noisy elements from within the wavelet domain. Their approach to image compression managed to achieve a level of performance, that surpassed the results produced by samples that were initially filtered in the spatial domain prior to image coding.

Starch and Murtagh [50] also employ wavelet compression to suppress noise from a digital image. The wavelet transform is used to convert the digital sample to the wavelet domain, where residual coefficients are analysed for sufficiently large structures at each scale. Collectively, these key-structures form the final restored image, with the remaining undesirable coefficients representing the original noisy data. In [51], Starch et al, continue to demonstrate the effectiveness of noise removal through image compression. Using an alternative multi-scale approach (pyramidal median transform), they successfully remove the all-pervasive noise from a selection of astronomical images.

The first order statistical pattern recognition technique [5, 52, 53, 54] represents an elementary approach to extracting a measurable feature from a digital image.

Weeks [55] uses numerous examples to highlight the one-dimensional histogram's system of grey tone analysis and object segmentation. He examines the histogram spread to detect and remove noise, and to isolate key objects within a scene using a basic thresholding technique. Weeks, then details the central moments that are utilised in an attempt to extract useful information regarding grey tone distribution. Pitas [28] also refers to the central moments and provides an example of image segmentation. Similarly, Gonzalez and Woods [29] offer further examples and remark on the histogram's limitation as a tool for image enhancement. Van der Heijden [26] uses the central moments to distinguish between the components that combine to form a blood cell and to isolate them from the background scene. In [56], Malagnini et al select the optimum set of training and test samples from an initial set of 450 images of galaxies, stars and background noise using the one-dimensional histogram.

The second order statistical pattern recognition technique [28, 29, 52, 57, 58] enabled a detailed analysis of the steel crystal structures, rather than simply offering a basic description of grey tone existence.

Haralick et al [3] applied the grey level co-occurrence matrix (GLCM) to three different types of natural texture data sets, and achieved high rates of classification for all three. They identified 89% of 100 sandstone test samples (also [59]), 82% of the eight category aerial photographic data set (similar [60]), and finally, 83% of the 310 images, seven category satellite imagery (also [61]), as opposed to 74%-77% using spectral

classification methods. Terzopoulos and Zucker [62] use the GLCM feature to identify the genetic disease Osteogenesis Imperfecta using digital images of fibroblast cultures. They claim that the GLCM, in conjunction with generalised co-occurrence matrices (GCM), outperforms a pre-selected team of six experts, with a classification rate of 90% and 69% respectively. Walker et al [63] combine the grey level co-occurrence matrix with Markov Random Field image features to classify cervical cells into normal and abnormal groups. Cox et al [64] use the second order feature to detect cancerous lung nodules and manage to classify at best 96% of the nodules and 93% of the tissue. Siew et al [65] assess carpet wear appearance using a number of texture analysis techniques including the grey level co-occurrence matrix. In [66], Conners and Harlow, perform a comparative study of texture measures, including the GLCM, grey level run length method (GLRLM), power spectral method (PSM) and the grey level difference matrix (GLDM). Applying the techniques to both synthetic textures and chest radiographic images focussing on the right lung, they concluded that the grey level co-occurrence matrix produced the most encouraging set of results, with overall classification rates of 90% and above. Finally, Van Gool et al [67] provide a detailed review of statistical and structural approaches to texture analysis, including a comprehensive report of the results obtained from their application including the GLCM feature.

The genetic algorithm (GA) [4, 52, 68, 69] and K-nearest neighbour (Knn) [5, 26, 52, 54] approaches to image classification were used to evaluate the discriminatory performance of the second order feature measurements.

Dudani [70] uses a weighted K-nearest neighbour (WKNN) scheme to classify 3000 test samples using 150 training images. Punch et al [71] advance the initial studies of

Siedlecki and Sklansky [72] to produce a hybridised GA and Knn classifier. The classifier assigns a particular GA weighting value to each feature descriptor, based on its ability to classify known samples. The result is a warping of the Knn space that is sensitive to those features that are important for the discrimination task. Applying the hybrid approach to real-world data, in the form of 300 training images and 60 test images, divided into three classes of plant soil, produced a 25% improvement over the standard Knn technique. Raymer et al [73] employ the same algorithm to predict water-mediated and polar ligand interactions in protein structures to an accuracy of 90%. Kelly and Davis [74] improve upon their weighted K-nearest neighbour (WKNN) approach in [75], and propose a rotated, WKNN scheme using the genetic algorithm search procedure. The technique rotates the current data point within the feature space to overcome any misclassification through its orientation with respect to the surrounding cluster measurements. The GA generates both the necessary rotation and element weighting values based on the initial training set. Operating on two artificial data sets, the system classified 96% and 87% respectively.

The rank-conditioned morphological analysis and template optimisation routines of Chapter 5, offered an alternative feature set that defined image texture in terms of surface interaction and pixel containment within a basic geometric structure.

Rees and Jones introduce the rank-conditioned (R-C) morphological transform in [76] and demonstrate its ability to extract both concave and convex image features in addition to enhancing the wound contours within an ulcerated region. In [77], they provide further evidence of its capability as a procedure for image enhancement, and include an example of texture separation. Rees et al [78] promote the R-C transform as a feature extraction

and analysis technique, that offers an accurate representation of the image surface and defines the interaction of the underlying texture components. Further examples and applications are presented in [6].

Research into producing the optimum template for morphological analysis remains strictly limited to the domain of image enhancement and noise removal rather than texture classification.

Soille [79] outlines the various types of structuring elements (SEs [79]) as pairs of points, adaptive SEs such as those proposed by Salembier [80], composite SEs, elementary symmetric SEs, and finally, non-symmetric SEs. Salembier [80] introduces the concept of adapting the shape of the analysing template in accordance to the desired output signal, and applies the procedure to a number of samples with some success. He tailors the shape and size of the template to the required image processing task including, noise removal, shape detection and image restoration through the removal of various types of image corruption. Dougherty and Loce [81] produce optimal structuring element libraries and examine noise corrupted binary images through a restricted observation window to further refine the filtering operation again for noise removal and image restoration. In [82], they expand the idea and define various filtering methodologies that specify the optimum filter according to the level of noise within the source image. Ehrhardt [83] and Harvey and Marshall [84] propose a method for morphological filter design using the genetic algorithm (also [85]) and demonstrate the resultant filter's ability to suppress noise.

2.5 Interim Conclusion

The literature survey has revealed only a limited amount of research material concerning the processes involved in successful galvanneal steel image feature extraction and classification. Nevertheless, the image processing procedures detailed in the previous section have been shown to be capable performers in the field of natural texture analysis. They are therefore deemed as being techniques that are suitable for galvanneal surface inspection, based on the work outlined in the referenced publications.

Chapter 3

Test Data Preparation

3.1 Introduction

Organisation and preparation of the source material prior to classification is a process that is as important as the selection and implementation of the feature extraction and analysis procedures. With any form of classification system it is essential that the test data be organised into well-defined groups. This requires a clear understanding of the parameters that define one particular class from another.

Difficulties are associated with image samples badly prepared for examination. Variations in image size and differences in grey-scale depth within and between each group, contributes to poor class assignment decisions as the applied technique has the potential to extract information which varies from sample to sample, offering no fixed practical in-class feature standard. Equally as important is the process of predetermining the analysis boundary region that defines the area of inspection. Continuity between each image ensures that the utilised approach behaves in a predictable fashion and produces a useful feature set accurately representing the source material. The classification problem may

then continue with the knowledge that the image sample representatives can be utilised with some confidence.

The above points may be considered obvious. Nevertheless, they are worth reiterating since they form important aspects of any classification system utilising image processing techniques as its foundation.

3.2 Image Properties

The galvanneal steel images possess an 8-bit grey-scale range and a resolution of 499 pixels by 386 pixels. Three sets of digital test samples have been extracted from actual steel-strip coils at varying levels of magnification, x50, x200 and x500. Figures 3.1a, 3.1b and 3.1c present the three fundamental types of annealing (under-alloyed, optimum-alloyed and over-alloyed), at the specified sampling resolution. Appendix 1 provides a small collection of images from each x500 class to demonstrate the variations in quality and crystal structure, and to highlight the problem of weak class-defining boundaries at the highest level of magnification.

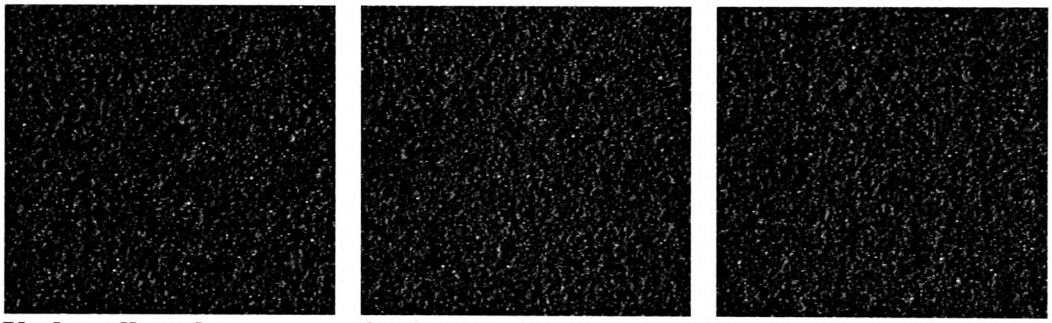
Analysis of the following examples in ascending order of magnification (Figure 3.1), it is apparent that the depth of field blur prevalent at the highest magnification level introduces an unavoidable trade off between the crystal size and image quality.

The x50 samples offer the clearest set of images, since the magnification level is well within the Leco image capture system's specification. However, there is no indication of any fine crystal detail or class-unique structural formations that would facilitate the distinction between the three annealing groups. Basing the classifier on a collection of

images that can not be easily separated increases the pressure placed on the feature extraction and analysis techniques to attempt to extract meaningful measurements from nondescript data where no unique characteristics exist. Therefore, the x50 sample set is considered to be too limited and unreliable as a potential source of classifiable galvanneal steel representatives. Consequently, the results extracted from the x50 images have not been documented.

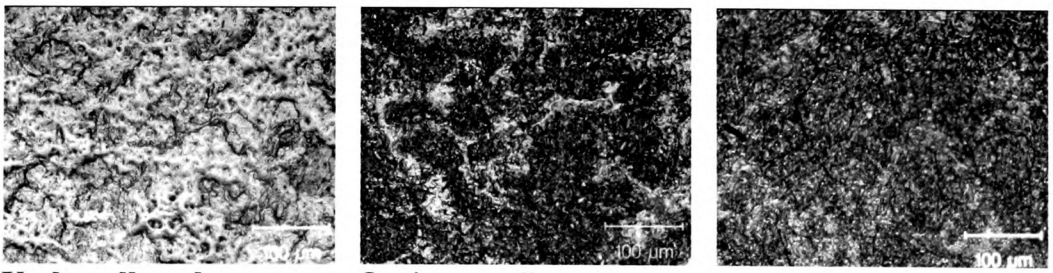
Figure 3.1: Surface view of the galvanneal steel coating at varying levels of magnification, sub-divided into annealing phase classes

(a) x50



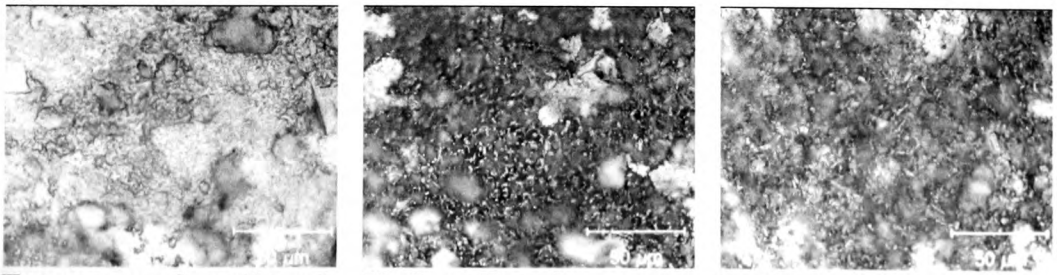
Under-alloyed **Optimum-alloyed** **Over-alloyed**

(b) x200

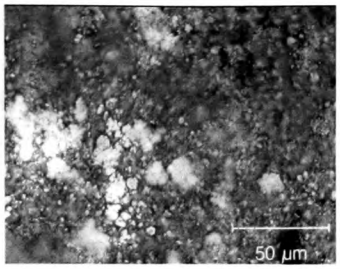


Under-alloyed **Optimum-alloyed** **Over-alloyed**

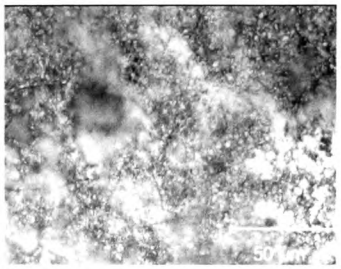
(c) x500



Extreme under-alloyed **Under-alloyed** **Optimum-alloyed**



Over-alloyed



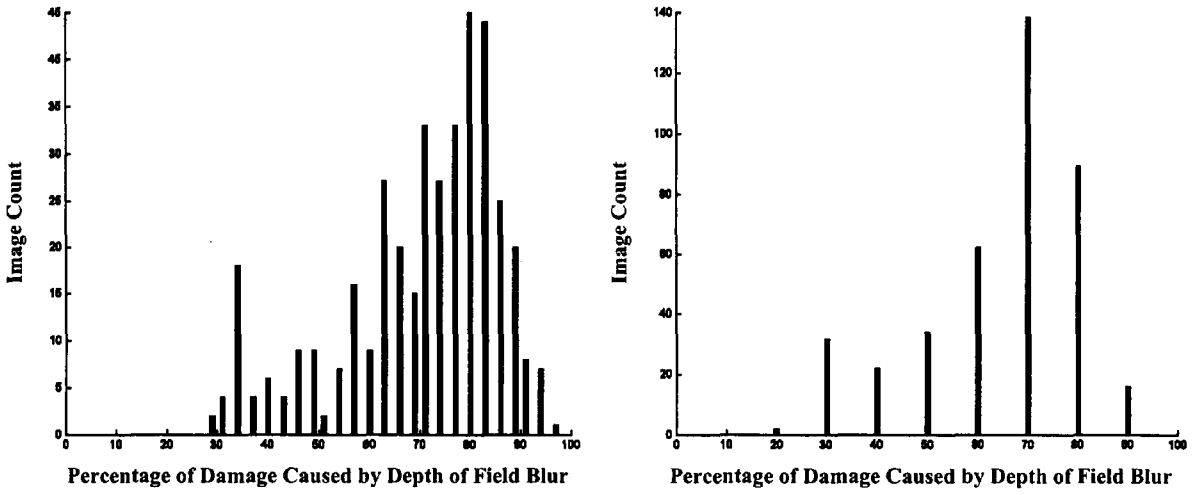
Extreme over-alloyed

The amount of detail greatly increases as the magnification level is raised to x200. It increases to such a degree, that it is now possible to pinpoint the under-alloyed strip from the remaining two classes. Separating the over-alloyed from the optimum state still proves to be a difficult task. The severe intensity differences that distinguished the under-alloyed image from the alternative groups is by no means as distinct. Nevertheless, surface flaws in the over-alloyed image are clearly detectable in the form of micro cracks and weak crystal formations, and offer potential targets for the galvanneal classifier. Only the low image count and class count of 120 and 3 respectively, prevent the x200 series from taking the position as the main training and testing data set. However, Chapter 6 uses the collection to verify some important points raised in Chapters 4 and 5.

As the optical lens approaches its magnification limit of x500, the image capture system introduces a problem that is not present in the previous examples. In almost every case within the x500 collection up to 97% of the crystal data is affected by a lack of definition, as the optical analysis system is unable to focus clearly on the steel-strip segment at the maximum resolution. Figure 3.2 presents the histogram and accumulative histogram defining the extent of the damage caused by the depth of field blur. Most images are affected to a significant degree, around the 70%-80% mark. This reduces the utility of the entire sample set by restricting the view of any surface defects (clearer in the x200

images) and crystal formations. Nevertheless, initial experiments will focus primarily on the x500 samples since the data set consists of a greater number of images and classes when compared with the alternative collections, 395 and 5 respectively.

Figure 3.2: Histogram and accumulative histogram outlining the percentage of x500 image damage caused by depth of field blur



(a) Histogram (b) Accumulative histogram

3.3 Image and Class Organisation for Texture Analysis and Classification

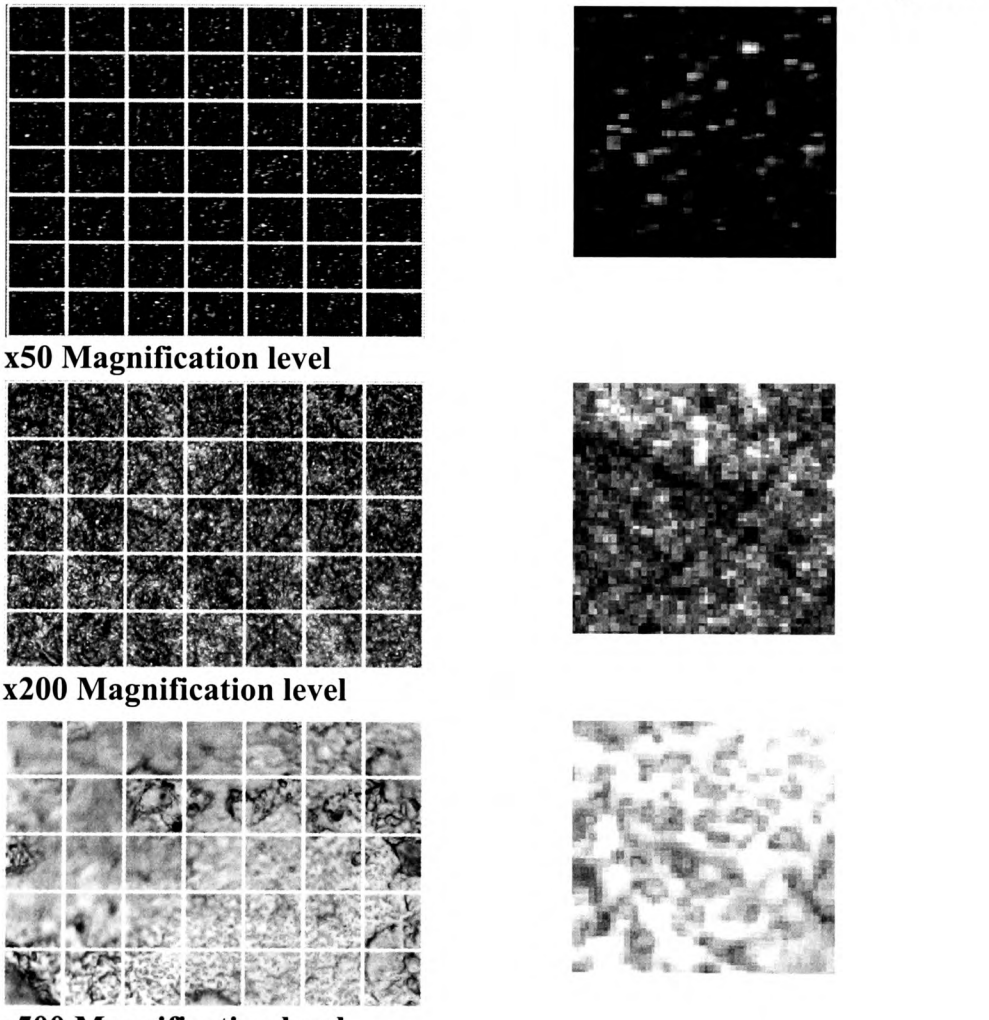
For the wavelet transform and the first order statistical technique (Chapter 4), the galvanneal steel images were divided into 64^2 pixels [61, 86, 87, 88] (Figure 3.3). Organising the samples into smaller regions addressed two main problems.

Firstly, with 35-49 segments per image, the amount of digital test data available is significantly increased. Even though the use of extracted portions from a single sample does not represent the full range of texture variations within a class, it still provides a useful testing and training collection to analyse at the early stages. Similarly, the amount

of fine level textural detail present within the 64^2 pixel segments is sufficient enough for the purpose of evaluating the performance of the first set of techniques, particularly where the extracted block is only barely affected by noise.

Figure 3.3: Galvanneal samples at various levels of magnification divided into 64^2 pixels

(a) Complete image **(b) Enlarged image segment**



x50 Magnification level

x200 Magnification level

x500 Magnification level

Secondly, the quantity of information that the applied techniques would have to process is dramatically reduced when compared to the amount stored within the original dimensions of 499 x 386. This is especially beneficial to the wavelet process, which required

approximately 7 minutes to analyse each segment under the interpreted environment of Matlab.

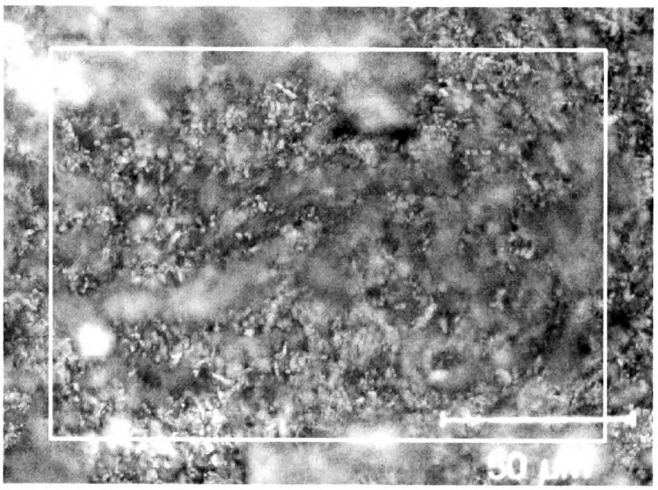
Reducing an image into 64^2 pixel regions also offered the possibility of analysing the sensitivity of the first order statistical approach to variations in the capture lighting condition. By traversing the image and examining the first order measurements extracted from each region, subtle changes in intensity over the image surface can be detected and any unfavourable effects on the feature duly noted. For real, live classification problems, using the smaller sized training samples for testing, provides a simple and effective method to assess the suitability and utility of any feature extraction and analysis procedure that relies directly on the grey-tone value of an individual pixel.

The remaining feature extraction techniques of Chapters 5 and 6 required some form of fixed boundary within the outer perimeter of each galvanneal steel image to prevent any unfavourable boundary conditions. Including the examination of data that does not form part of the original source image i.e. data residue stored within the computer system's memory.

For example, the rank-conditioned morphological procedures of Chapters 5 and 6 operated exclusively on the entire image rather than the smaller 64^2 pixel segments. Their efficient rate of operation and a more relaxed attitude to variations in lighting across the whole image eliminated the need to restrict the size of the referenced sample. Nevertheless, the operating nature of the morphological process made it essential that a restrictive boundary be utilised. As a measure to reduce the ability of the blurred regions

to affect the accuracy of the final feature, the border size remained fixed at 427 x 314 leaving a breathing space of 36 pixels around the entire perimeter (Figure 3.4).

Figure 3.4: Restrictive image border employed to eliminate the risk of unfavourable boundary conditions when performing the rank-conditioned morphological transform (Chapters 5 and 6)



The majority of the applied feature extraction and analysis procedures of Chapters 4 and 5 operate on the complete set of 395, x500 galvanneal steel images. No distinction is made between an organised training and testing set, the collection is simply divided into classes based upon the percentage of iron present in the steel substrate. The three fundamental annealing groups of the x50 and x200 samples (under-alloyed, optimum-alloyed, over-alloyed) have been further sub-divided into five classes with the addition of an extreme under and over-alloyed phase. Table 3.1 defines the five possible annealing phases and the level of iron within the protective coating. For the genetic algorithm and K nearest-neighbour classifiers of Chapter 4, the classes have been referred to using pre-defined labels (columns 3 and 4). The labels refer to smaller subsets extracted from the original classes to give around 30 images for each training and testing set and an overall sample

count of 300. This prevents any advantage/disadvantage gained from one particular class dominating the alternatives by either intensifying the density of its cluster, or by increasing its spread and encroaching upon the boundaries of the remaining groups. Maintaining a common sample count not only preserves the impartial status of the applied classifiers but also ensures that the initial solution space does not favour any one type.

Table 3.1: Outline of x500 annealing class organisation for experimental training and testing

ANNEALING TYPE	APPROXIMATE IRON %	TRAINING CLASS LABEL	TESTING CLASS LABEL
Extreme under-alloyed	5.38	B	A
Under-alloyed	8.42	D	C
Optimum-alloyed	10.79	F	E
Over-alloyed	12.08	H	G
Extreme over-alloyed	14.37	J	I

Where appropriate, the preparation, organisation and labels remain constant for the alternative x200 samples and the various texture classes of the Brodatz collection (introduced in Chapter 6). Any variations in class analysis and image ordering are clearly outlined prior to the stages of experimental examination and classification.

3.4 Interim Conclusion

The three sets of galvanneal steel images offer a trade off between crystal clarity and crystal detail. Whilst an increase in the level of magnification results in an improvement in the amount of fine structural detail available, the quality of the texture remains variable and lacks the focus observed at the lower resolution levels. Nevertheless, the choice of a central training collection is limited to the x500 image set due to the presence of a higher number of samples and an increased class-count. Even though x200 images offer the best

balance between clarity and structural detail their use is restricted to specialised operations and experiments. Chapter 6 sees the introduction of the x200 set to evaluate the drop in feature accuracy caused by the blurred regions and examines the effect on the level of classification. Similarly, the Brodatz set (Chapter 6) is examined to assess the performance of the investigated techniques when applied to alternative forms of natural texture.

Data preparation and class organisation is a significant issue that requires addressing in any classification problem. Here, the subject has been approached in a reasonable manner by organising the samples into natural groups based on the iron content within the steel substrate. Additionally, the galvanneal images have been prepared in accordance with the operating restrictions of each applied technique. Through the introduction of a restrictive image boundary or by dividing each sample into smaller sections, the data is prepared and arranged in such a way, as to limit the negative effects that poor class and image organisation may have on classification performance. For example, if each class boundary is only weakly defined, then the classifier may be wrongly blamed for making poor decisions. This may force an unnecessary change in direction over the combination of the applied feature representative and classification system, and reduce the likelihood of attaining the optimum result.

Chapter 4

Galvanneal Steel Image Analysis using Established Image Processing Techniques

4.1 Introduction

This chapter reviews the preliminary techniques applied to the steel images in an attempt to classify the galvanneal samples at an early stage using well-established methods. The introduction and description of each procedure has been intentionally restricted given the amount of literature available on each topic. Key references are cited throughout the text to supplement those referenced in the second chapter.

The lack of available research material into a possible connection between the coating type of a steel sample and its optical representative, has led to the application of established techniques and the introduction of a "back-to-basics" approach to texture analysis and crystal detection. By commencing the study at a position where the wealth of material available is well known and well tested, some level of expectation is provided on how the applied techniques will perform. It would also offer a comparative base on which to test the performance of a potential galvanneal steel classifier, against other relevant, established or newly developed routines.

Summary of applied techniques:

(i) Wavelet transform - used to compress data for the removal of elements between the galvanneal steel crystals. It was unable to isolate any class specific features and offered no improvement over the quality of the standard images. Led to the following technique of extracting and exploiting the differences in pixel intensity between each annealing class.

(ii) First order statistical technique - analyse image intensity. The histogram procedure was too sensitive to the value of an individual pixel and produced results that failed to indicate any form of in-class image homogeneity. This influenced the choice for the next method of examining the crystal structures through the interaction of pixel-pairs.

(iii) Second order statistical technique - evaluate structural information. The grey level co-occurrence matrix is less affected by the value of an individual pixel. It attempts to model the surface texture based on the variations in pixel-pair intensity over the entire image. The technique produced a promising set of a priori results but failed to produce tight, isolated clusters to support classification without any prior knowledge. The results generated by the co-occurrence matrix were analysed and categorised using the following classifiers.

(iv) Genetic algorithms - examine how the second order results perform in a non-a priori environment. Genetic algorithms were used to reduce cluster variance and

improve group separation by implementing an element-weighting scheme. This method was unable to rectify the problem of cross-cluster contamination.

(v) K-nearest neighbour - evaluation of the second order results using an alternative classification routine. The K-nearest neighbour classifier defines the class of an incoming data point based on its K nearest neighbours. This routine was incapable of producing tight clusters and was affected by the order in which the test data was analysed.

4.2 Wavelet Transform

4.2.1 Introduction

The wavelet transform [51, 89, 90, 91] is a powerful signal decomposition and synthesis technique that has a wide range of applications [92, 93]. These range from data analysis [94], image restoration and enhancement [50, 95, 96, 97, 98], feature extraction [99], multiresolution analysis [1, 100, 101] and data compression [102, 103]. It uses a family of orthonormal basis obtained through the translation and dilation of a central function known as the "Mother wavelet". The transform offers a more appropriate method of analysing non-continuous signals with variations in frequency behaviour than the sine and cosine basis functions of the Fourier transform, which by definition are non-local and extend to infinity [104]. Conversely, the approximating functions of the wavelet transform are localised in both space and frequency, properties that provide an increased level of accuracy in data manipulation and the analysis of signals that contain short high frequency components, and extended low frequency components [105].

4.2.2 Wavelet Analysis

The two-dimensional continuous wavelet transform [1] is presented as:

$$W(a,b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x,y)\Psi_{a,b}(x,y)dxdy$$

Giving the discrete transform,

$$W(a,b) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(x,y)\,\Psi_{a,b}(x,y)$$

For an $M \times N$ image matrix, where $\Psi_{a,b}$ is the wavelet function and a and b are the dilation and translation scaling factors.

Initially focussing on the discrete orthonormal basis as outlined by Daubechies [1, 51, 106], where every basis function is obtained from the central mother wavelet $\Psi_{j,k}$ by shifting and scaling:

Gives,

$$\Psi_{j,n}(x) = 2^{\frac{-j}{2}}\Psi(2^{-j}x - n)$$

Where j represents the wavelet's width, and location index n gives its position.

Spanning the data domain at different resolutions is achieved by using the mother wavelet in the scaling equation [44]:

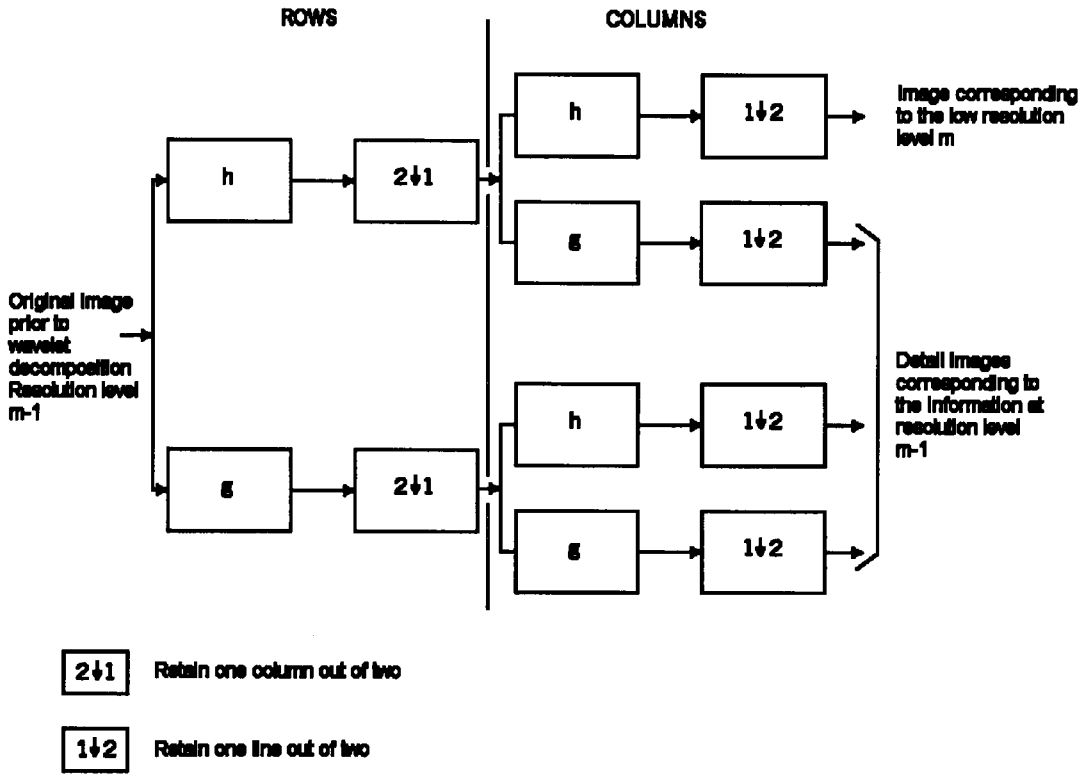
$$W(x) = \sum_{k=-1}^{N-2} (-1)^k C_{k+1} \Psi(2x + k)$$

Where $W(x)$ is the scaling function for the mother wavelet Ψ and C_k are the wavelet coefficients. These coefficients can be thought of as an image filter [104] which are further placed into a transformation matrix [44]. The components are ordered in such a way that when the structure is applied to the raw image data it is able to both smooth and enhance the present information. This is achieved by arranging the matrix into a quadrature mirror filter pair format [107] where the odd rows contain the low-pass ordering of the wavelet coefficients for the smoothing operation, and the even rows contain the complementary high-pass data to enhance the image's detail.

The matrix is applied in a hierarchical, decompositional method known as the pyramidal algorithm [51, 108], as shown in Figure 4.1. It first acts upon the entire full-length image vector, which is then smoothed and reduced to half its size. The process of applying the matrix followed by smoothing and halving is repeated until a trivial amount of smooth data remains, at which time the image decomposition process comes to an end. The final output of the discrete orthonormal wavelet transform, is composed of the remaining smooth and all the directional detail components, resulting in the original image being reproduced in the wavelet domain. Once the image has been reproduced, a number of data operations can be carried out on the corresponding wavelet coefficients. For example, by truncating these components below a predefined threshold value, non-essential data can

be removed. Reconstructing the image from the remaining coefficients produces a compressed equivalent, which in certain cases is indistinguishable from the original.

Figure 4.1: One stage, two-dimensional image decomposition algorithm [108]



4.2.3 Experiments and Discussion

Using the Uvi-Wave toolbox for Matlab, the following steps were carried out in order to compress the supplied image set:

(i) Import raw image data as a 2-D matrix.

(ii) Generate low and high pass wavelet analysis filters.

(iii) Select a two-dimensional wavelet packet basis [107] using the pruning algorithm and additive cost measurement (best basis/tree algorithm) [109].

(iv) Convert data to wavelet domain.

(v) Drop specified percentage of coefficients.

(vi) Reconstruct spatial data by calculating the inverse wavelet transform.

By carrying out this lossy data compression technique and removing a progressively greater percentage of the smaller wavelet coefficients, it was theorised that some underlying feature would become prevalent in each of the five classes (x500 image set). It was expected that the filtering operation would remove the pixel pathways that interconnected each class-specific texel island. These islands could then be extracted and measured to help form the basis of a classification feature set.

64^2 pixel segments were extracted from each of the 395 images to measure "sub-image" differences in addition to in-class and between class differences. Each segment was passed into the previously defined wavelet compression routine using the Uvi-Wave toolbox and the following experiments were carried out:

- Analysis and class comparison of the compressed image coefficients in the wavelet domain, to find out if there is close correlation within each of the five classes and to see if the groups are clearly separable in the wavelet domain.

- Accumulative histogram on the above compressed coefficients, to highlight any significant group trends.

- Small and large co-efficient thresholding to carry out simple key-trend analysis.

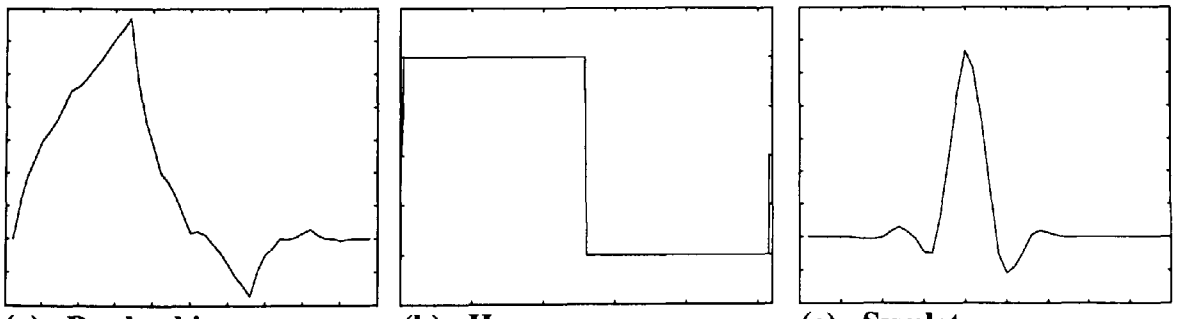
- Pre-processing the raw image data to see if crystal structures could be further enhanced prior to the application of the wavelet transform.

- The use of alternative wavelet analysis filters (Daubechies, Haar and Symlet) (Figure 4.2) to examine whether there would be any perceived differences in the filtered representation of the original image.

- Retaining a progressively lower percentage of wavelet coefficients (90%, 50%, 20%, 10%, 1%) to observe the effect of wavelet compression and to pinpoint the stage at which class defining crystal structures become the dominant feature.

- Analysis and comparison of the reconstructed image data after co-efficient reduction, to examine the effect on the remaining texel formations.

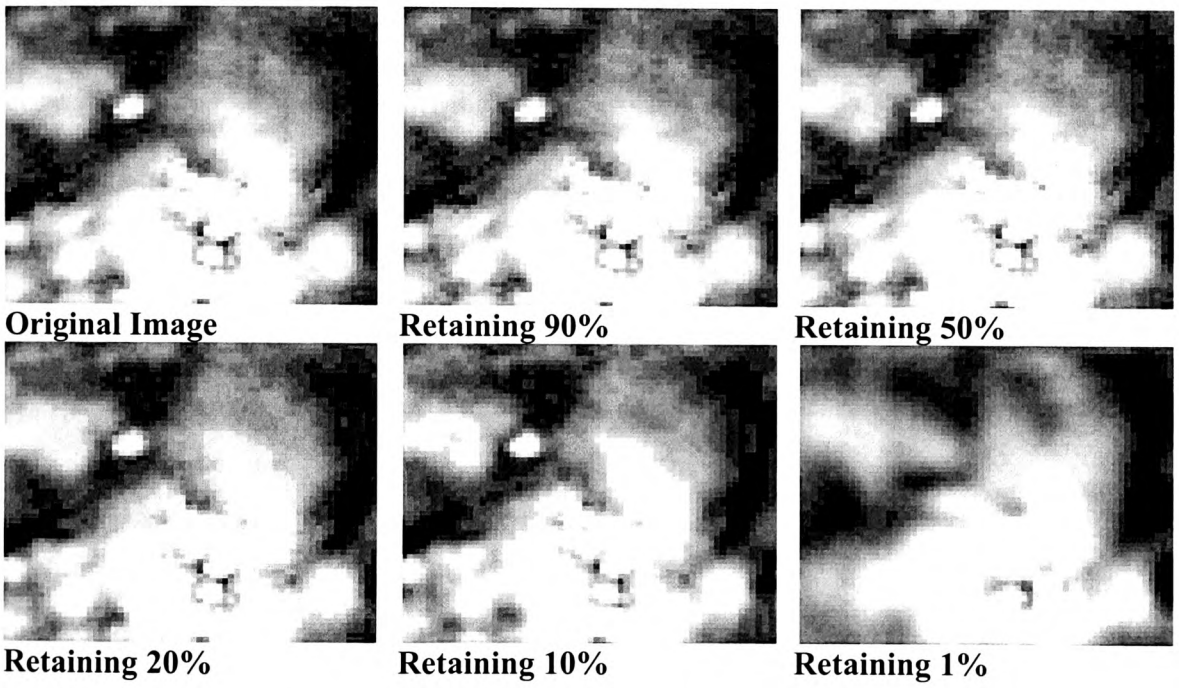
Figure 4.2: Selection of applied wavelet functions



(a) Daubechies **(b) Haar** **(c) Symlet**

Even for high losses, the filtered and original images remained similar. At a retention level of around 20-10%, the filtered images began to lose detail as fewer coefficients were used to retain the areas of higher intensity (Figure 4.3). The net result was an image with fine detail corrupted (in addition to the depth of field blur) by the wavelet filtering routine. For the purpose of classification, no improvement was found.

Figure 4.3: The effect of the wavelet transform on a single galvanneal image segment at varying levels of compression (64^2 pixel segments)



At no stage of the compression process did any class defining crystal structures present themselves. This is due to the fact that the crystal features are typically made up from low intensity pixels (with the exception of the extreme under-alloyed class and to a certain extent the extreme over-alloyed class). The wavelet compression process tends to consider these darker regions as non-essential background information due to the intrinsic operation of the quadrature mirror filter and the coefficient removal scheme. By removing

the lower frequency elements, the system was actively targeting the lower intensity regions. Where the areas of high intensity were reconstructed at a similar resolution level to the original, the "background information" (the very feature the compression routine was expected to enhance) was adversely affected by the low pass analysis filter, resulting in a blurring of the crystal data. Essentially, the wavelet transform served only to aggravate the already significant problem of neighbourhood grey-scale averaging, an inherent problem resulting from limitations with the optical analysis system.

The application of alternative basis functions was an attempt to discover whether filters of varying symmetry [1] would retain alternative texel structures within the re-constructed compressed images. Analysis of the re-constructed data revealed that all three functions generated images of similar quality at each filtering level. The only perceivable difference came from the results generated by the most compact and symmetric of the three basis, the Haar wavelet [1]. This produced a slightly less accurate filtered image as the coefficients were unable to resolve the fine detail of the steel texture, giving the final compressed product a washed out look. I.e. the fine local variations in pixel intensity have been replaced with blocks of similar grey-tone elements, causing a noticeable loss in textural detail. Similarly, any attempt at pre-processing in the shape of image sharpening served only to add noise to the non-blurred regions, whilst further deteriorating crystal clarity within the already obscure areas. Histogram equalisation [26, 53] adversely affected the image by boosting the darker regions around the high intensity areas, rather than enhancing potential classification features. This resulted in the wavelet transform losing yet more class specific material while retaining the non-essential data.

Measurement of the converted image coefficients directly using the accumulative histogram technique displayed minor similarities within each class but proved to be unreliable. Significant differences were observed within a large percentage of the 64^2 pixel sub-images, with a considerable amount of crossover results between the three basic classes of annealing (under-alloyed, over-alloyed and the optimum-alloyed states).

4.2.4 Conclusion

The wavelet transform failed to offer a viable solution, through pre and post data compression using direct coefficient or spatial measurements. Its impressive ability to represent an image at such a high standard using only a fraction of the available information actually prevented the isolation of key class features. The transform attempts to elegantly remove background information, which would not obviously be detected by the human eye. This graceful form of filtering not only failed to perform the brutal form of pixel removal required to segregate the class-defining crystals from the blurred areas, but also reduced the clarity of the unaffected texture primitives.

Statistical measurements have been extracted from the wavelet-compressed images and the results are presented in the next section.

4.3 First Order Statistical Pattern Recognition Technique - Probability Density Function

4.3.1 Introduction

The probability density function (PDF) [28, 29] is a well-known spatial feature extraction technique. It provides baseline data that offers the ability to compare the efficiency of the applied methods, in relation both to their actual performance and to their performance relative to the work of other researchers in the field of optical texture analysis and classification.

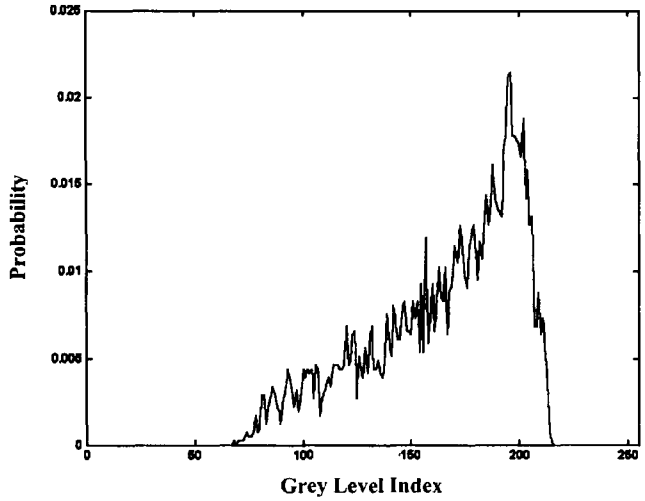
Consider the grey-scale image intensities $f(x,y)$ as being a set of random variables r that represent the probability of the grey-tones existing in the interval $(x, x+dx)$, written as, $\text{Prob}(x \le r \le x+dx)$. It is possible to generate a measurable feature that conveys important information about global grey-tone distribution. This feature is commonly used in image segmentation and classification [110, 111, 112] due to its information content and mathematical simplicity.

The PDF is estimated from the image histogram P_f, which represents grey-scale frequency and is given by the relation [28],

$$P_f(f_k) = \frac{n_k}{n} \qquad k = 0,1,...,G-1$$

Where $G \in \{0..Ng-1\}$, where Ng represents the number of discrete grey-scales and n_k is the total number of pixels that possess a grey-level intensity of k. Variable n depicts the total number of elements in the two-dimensional image matrix, and normalises the original histogram to give the PDF feature (Figure 4.4 shows a typical PDF).

Figure 4.4: Typical probability density function



4.3.2 Probability Density Function Analysis

The four central moments described by Levine [113] provide a collection of statistical measurements that offer direct analysis of this one-dimensional feature. The four moments mean, variance, skewness and kurtosis are given by,

$$\text{Mean} = \mu = \sum_{k=1}^{N} f_k p_f(f_k)$$

Where f_k represents the various image intensity levels and $p_f(f_k)$ the normalised image histogram. The arithmetic mean of a probability density function is defined as a measurement of the data's tendency to rest around the central area of the grey-scale axis.

It is considered to be one of the more familiar central tendency measurements available, but has the disadvantage of being sensitive to extremities, a problem that has a direct effect on all of the three subsequent moments.

$$\text{Variance} = \sigma^2 = \sum_{k=1}^{N} (f_k - \mu)^2 p_f(f_k)$$

The second moment variance, uses the results from the previous calculation to evaluate the dispersion of grey-level intensity around the central location and is otherwise known as the standard deviation. Larger values are generated for the PDF that displays a greater level of element dispersion.

$$\text{Skewness} = \mu_3 = \frac{1}{\sigma^3} \sum_{k=1}^{N} (f_k - \mu)^3 p_f(f_k)$$

Histogram skewness characterises the degree of asymmetry of pixel distribution around the mean point and is a value that possesses no measurement label to signify its type. If the histogram is symmetrical about the mean, then the skewness function returns a value of zero, if it's skewed left or right then the result is negative or positive respectively.

$$\text{Kurtosis} = \mu_4 = \frac{1}{4} \sum_{k=1}^{N} (f_k - \mu)^4 p_f(f_k) - 3$$

The measurement of kurtosis, the fourth moment about the mean, is also a purely descriptive type that represents the degree to which a distribution is flat or peaked. Where the larger values are produced for the distribution that can be regarded as above normal

(leptokurtic), smaller values for the flatter configuration (platykurtic) and mid-range values for the feature with a normal distribution (mesokurtic) [44].

Entropy is also defined in terms of an image histogram,

$$\text{Entropy} = V = -\sum_{k=1}^{N} p_f(f_k) \ln p_f(f_k)$$

The entropy of an image defines its structure in terms of the distribution of local grey-scale intensities. A uniform image will consistently return lower entropy values than the image with a random distribution of varying grey-tones.

4.3.3 Experiments and Discussion

The inability of the wavelet transform to isolate any well-defined crystal structures, led to an attempt to classify each sample based solely on grey-tone frequency. By generating the one-dimensional histogram from each galvanneal image and extracting the five predefined measurements, it was expected that this first order feature would generate similar in-class distributions, while providing enough scope to separate each annealing type.

Again, the 395 images were divided into 64^2 pixel segments to examine whether variations in lighting over the entire sample would have a detrimental effect on the final feature measurement. Histograms were generated for each image/segment prior to feature analysis and assessment. The areas covered include:

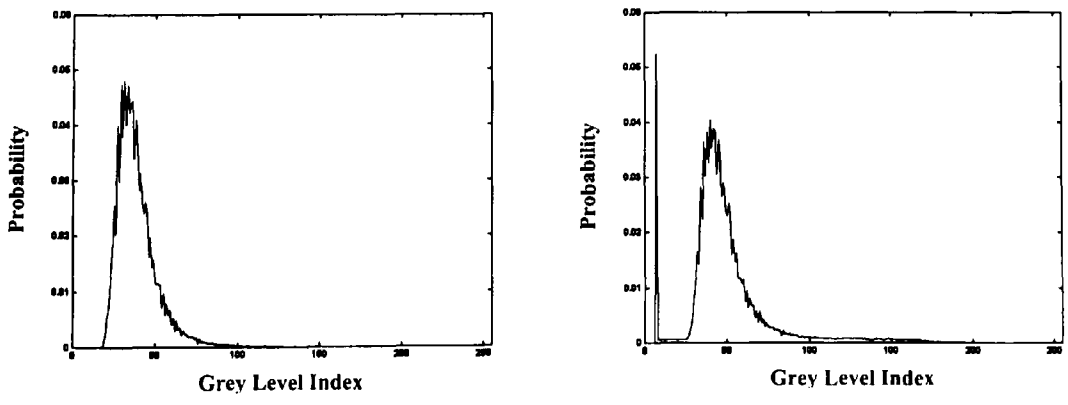
- Analysis and class comparison of the five measurements extracted from the raw image PDF. This would help to evaluate the effectiveness of a measurement based exclusively on intensity distribution rather than structural formation.

- The use of images in various annealing states, that have been acquired under different lighting conditions. This would show up any inadequacies in using a histogram feature for measurement in a real life, on-line situation.

- Pre-processing the raw image data using the wavelet transform, which would highlight any significant benefits in compressing the original image when carrying out first order feature measurement.

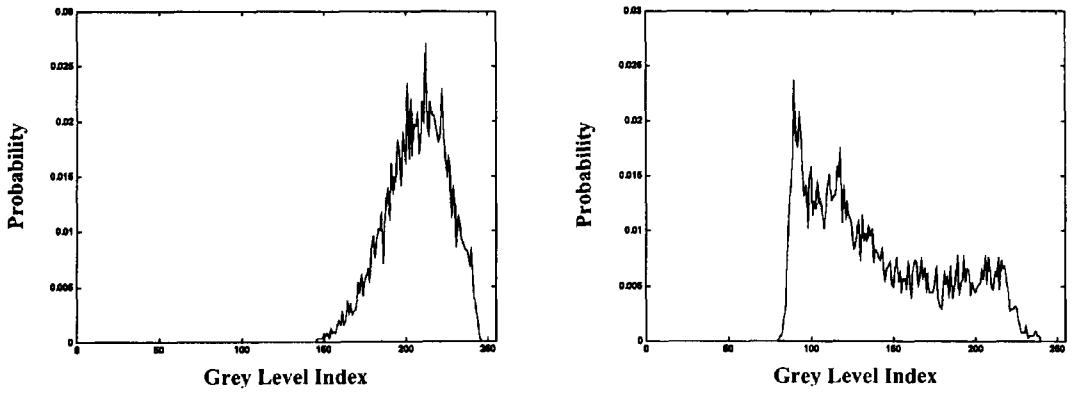
Histogram techniques are greatly affected by their reliance on the intensity level of a single pixel. This point was verified by analysing images of the same galvanneal type, which had been captured using alternative light sources or a varying angle of incidence. Figure 4.5 illustrates the distinct differences with two PDFs extracted from images of the same coating class, but captured in different environments.

Figure 4.5: Two PDFs extracted from the same annealing class, captured under different lighting conditions



The problem was also detected during the analysis and class comparison experimental stage, where the standard galvanneal images were split into the 64^2 pixel segments and their probability density functions measured (Figure 4.6). The subtle shading differences over the entire image had produced variations in the feature measurements, that were as widespread as those extracted from the images captured using different lighting filters. The great variations observed within each PDF calculation, further highlighted the histogram's sensitivity to the surrounding environment. By following Rosenfeld and Troy [40] and normalising each of the 395 images, the extracted histograms would possess approximately the same standard grey-level distributions and would overcome the differences in image shading variations. Whilst the data normalisation routine was successful, the results, when analysed, failed to reduce in-class differences or improve between class separation.

Figure 4.6: Two PDFs extracted from different sections of the same image



The images compressed using the wavelet transform were measured to ascertain whether the compression routine had a beneficial effect on intensity distribution by retaining only certain key elements. Results produced by the original wavelet transform (section 4.2.3), were imported into the first order algorithms. Using the same levels of compression (Retaining 90%, 50%, 20%, 10% and 1% of the coefficients), a new set of probability

density functions was generated and the four moments, in addition to the entropy measurement, were calculated.

Measuring the wavelet-compressed images confirmed its remarkable ability to accurately represent the original data using only a fraction of the original coefficients. Initially, the results for the lower compression levels mirrored those of the original images. But, as the retention level was reduced to 20%, the sub-image measurements showed a slight decrease in variation and therefore an increase in image correlation. At this low level of coefficient retention, the lossy compression routine removed a substantial amount of data, whilst retaining certain key elements that had helped to produce the similar histograms for each image segment. The danger with only a small amount of data remaining, is that the similar information present in the histogram may actually be the noise within the blurred regions, rather than important classifiable data. This premise was confirmed when the amount of between class crossover actually increased by an additional 70% as the retention level was progressively reduced. The high pass filter had only enhanced the detail of the few remaining high intensity elements and the low pass filter had further reduced the surrounding pixel definition.

4.3.4 Conclusion

The one-dimensional probability density function based on a standard, normalised or compressed image was found to be too sensitive to the surrounding lighting conditions and sampling environment. This sensitivity suggests that if the feature is unable to confidently classify images captured off-line, then it would certainly be unable to perform in the harsh, dusty environment of a steel strip galvannealing line.

4.4 Second Order Statistical Pattern Recognition Technique – Grey Level Co-occurrence Matrix

4.4.1 Introduction

The grey level co-occurrence matrix (GLCM) [58] can be considered as the natural approach to adopt upon completion of the first order histogram technique. This second order feature is a well-known statistical measurement tool, that describes the grey-level spatial interdependencies of the surface texture.

Ohanian and Dubes [114], have compared the classification performance of four very different types of textural features, including, Gabor filters, Markov random fields, fractal geometry and the grey level co-occurrence matrix, and concluded that the latter consistently returned higher recognition rates. Additionally, Tully et al [115] and Conners and Harlow [116], determined that the GLCM feature outperformed the grey level difference matrix (GLDM), spectral density function and the grey-level run length matrix, when describing texture parameters.

The co-occurrence approach defines the image in terms of the frequency at which a grey-scale pixel-pair exists within a predefined region. This provides the means to evaluate spatial organisation and more importantly assess the structure formations of tonal primitives, thus giving a detailed description of the image texture. Coarse textures are generally represented by only a slight change in distribution, whilst the finer textures produce a distribution that changes rapidly with distance [58]. The GLCM can be considered as the estimation of the second order joint conditional probability function,

$p(i, j \mid d, \theta)$, where the probability of going from grey-level i to j depends upon distance d and angle θ. These angles are formally quantized to 45 ° intervals and are typically measured at a distance of 1 or 2 pixels as suggested by Haralick et al [3].

The resultant feature p, is a square matrix, where the dimensions are determined by the number of discrete intensity levels present in the sample image. Each pixel-pair is used as a co-ordinate reference, where the corresponding feature points are incremented to register the frequency of their co-occurrence. The outcome of this process, is a measurable matrix that describes the grey-level inter-relationships of spatial components.

4.4.2 Grey Level Co-occurrence Measurements

Haralick et al [3], proposed a number of GLCM measurements that enable the extraction of distinguishing features useful for texture classification. Although these measurements contain essential detail about the structure of the image surface, it is often difficult to identify the specific texture characteristic being analysed [3, 114]. For this reason, a short definition accompanies each applied algorithm.

Using a normalised GLCM for each image, the following feature measurements were extracted,

$$\text{Maximum probability} = f_1 = \max(p(i, j))$$

This is a measure of the most common occurring grey-scale pair at an angle θ and distance d. It offers some indication of image homogeneity, where the homogeneous

image would return the higher values, since it would contain very few dominant grey-tone transitions.

$$\text{Entropy} = f_2 = -\sum_{i,j} p(i,j) \ln p(i,j)$$

Entropy defines the randomness of the surface texture, where the higher values reveal a greater level of randomness [3].

$$\text{Element} - \text{difference of order } m = f_3 = \sum_{i,j} |i-j|^m \; p(i,j)$$

This measures the local grey-level variations in an image. The descriptor is maximised when the image consists of a predominantly coarse surface. Moments $m = 2,3$ and 4 were used.

$$\text{Energy} = f_4 = \sum_{i,j} p(i,j)^2$$

Energy provides a definition of grey-tone transition. It yields higher results for the co-occurrence matrix that contain elements that possess values that range from very high to very low.

$$\text{Inverse Difference Moment} = f_5 = \sum_{i,j} \frac{1}{1+(i-j)^2} p(i,j)$$

The inverse difference moment (IDM) is another measurement that defines grey-tone transition and therefore local homogeneity. Results are minimised for coarse images, where the corresponding feature matrix contains coefficients that are primarily concentrated around the main diagonal.

$$\text{Shade} = f_6 = \sum_{i,j} (i + j - \mu_x - \mu_y)^3 \, p(i,j)$$

Where μ_x and μ_y are the mean values of p_x and p_y. This feature measures the degree of surface reflectance, whereby the higher value represents an image that is mainly comprised of high intensity pixels.

$$\text{Inertia} = f_7 = \sum_{i,j} (i - j)^2 \, p(i,j)$$

This texture descriptor measures the local grey-scale variations and is equivalent to the element-difference moment of order (EDMO), $m = 2$, in that it defines image coarseness [40].

$$\text{Prominence} = f_8 = \sum_{i,j} (i + j - \mu_x - \mu_y)^4 \, p(i,j)$$

This measurement, like shade, is affected by the degree of surface reflectance, where the maximum values are returned for the GLCM that consists of a low directional average and a large number of coefficients that represent the co-occurrence of high intensity pixel-pairs.

$$\text{Correlation} = f_9 = \sum_{i,j} \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} p(i, j)$$

Where σ_x and σ_y represent the standard deviation of p_x and p_y. Correlation measures the degree to which the rows and columns of a grey level co-occurrence feature resemble one another. Low results are generated from a matrix with large differences in element values.

Other matrix measurements have been defined [3, 58, 60, 117] but these offer no real advantage over the above algorithms when dealing with images formed from natural texture elements.

4.4.3 Experiments and Discussion

The analysis window was expanded from the initial 64^2 pixel segments to almost the size of the entire input image, giving an area of 497 x 384, restricted to prevent any unfavourable boundary errors. Local variations in texture formations throughout the image would not be accurately detected using the minimal window size, since the GLCM would not be effective at representing a wide spread of rarely occurring grey-tone pixels. The resultant matrix would be sparsely populated [64] and fail to offer a complete and accurate description of the galvanneal crystals.

Sub-image comparisons at this stage were not necessary. The co-occurrence feature defines spatial inter-dependencies rather than singular grey-tone existence and therefore is not as greatly affected by the value of a single pixel as the highly sensitive first-order technique.

The following experiments were carried out on the normalised co-occurrence matrices in order to establish a feature that would be suitable enough to differentiate between the predefined coating conditions:

- Generating the two-dimensional second order joint conditional probability density functions with the following displacement distances, d and angles, θ. This would highlight any directionality within the image texture and also the variation in distance may give an indication of coating crystal size:

  - $d = 1, \theta = 0^{\circ}$
  - $d = 1, \theta = 45^{\circ}$
  - $d = 1, \theta = 90^{\circ}$
  - $d = 1, \theta = 135^{\circ}$

  - $d = 2, \theta = 0^{\circ}$
  - $d = 3, \theta = 0^{\circ}$
  - $d = 5, \theta = 0^{\circ}$
  - $d = 10, \theta = 0^{\circ}$

- Pre-processing in the form of grey-scale reduction, in an attempt to emphasise potential key-texel structures. The following distances, d and angles θ, were used on images with 16 discrete grey-levels:

- $d = 1, \theta = 0^{\circ}$
- $d = 1, \theta = 45^{\circ}$
- $d = 1, \theta = 90^{\circ}$
- $d = 1, \theta = 135^{\circ}$

- Variation and mean measurements based upon the previous results in a final attempt to successfully classify each annealing type.

- The separation of each annealing class into smaller clusters in order to assess in-group variations.

At this stage varying the angle of pixel co-occurrence proved to be unnecessary, leading to the conclusion that the two-dimensional galvanneal data is asymmetrical. Furthermore, the majority of steel samples failed to exhibit any clear form of regular, periodic crystal pattern. This made it difficult to select an optimum displacement value [88, 118] that would best discriminate between the surface structures and the background noise. Varying d only produced a slight shift in scale whilst the variations of each measurement spread remained almost constant. The GLCM feature was still being affected by the widespread crystal blur irrespective of the value set for the distance parameter. Increasing its value to beyond the recommended 1 or 2 [3] to decrease the probability of the second reference pixel residing within a localised blurred region produced a matrix that was no longer concerned with the local variations in crystal structure, especially at distances of around 10 and above. Nevertheless, with the support of a priori knowledge the classification results indicated that crystal size detection was not an essential requirement

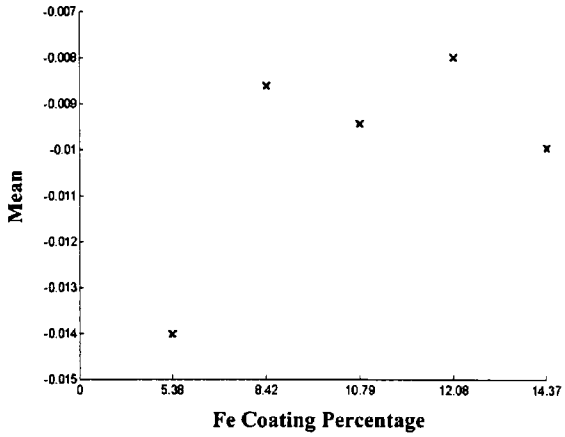
in differentiating between each coating class type. Figure 4.7 represent a selection of the results based upon the mean and variance measurements of the previously defined texture descriptors. The grey level co-occurrence features were calculated using a displacement distance of 1 pixel, an angle of 0°. The results indicate that using a minimum distance classifier [54, 119] and a combination of the available feature measurements, it is possible to confidently confirm the predefined coating condition of each of the 395 images.

To verify the successful distinction between the three main annealing levels, each class was split into equal groups and further analysed using the same GLCM parameters and measurements. Considering the rather poor image quality and low image count within each cluster, the a priori results indicated a high degree of similarity within each annealing class; but with the exception of the shade, prominence and correlation measurements, the level of crossover was unacceptably high. This indicated that using these poor quality images, the fundamental class differences are based on both variations in surface reflectance and on a less significant level, structure formation.

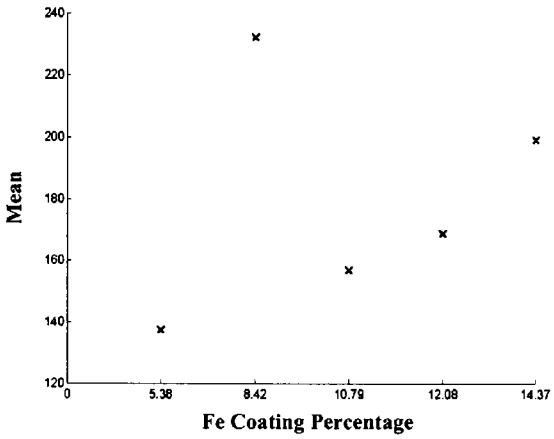
Elimination of the less common grey-tones from each galvanneal image emphasised the significance of the more common intensities. The original 8-bit data set was imported into Paint Shop Pro and restricted to a 4-bit, "Optimised Median Cut" [120, 121] palette. Grey-scale reduction returned surprisingly similar results to the original 8-bit images despite a reduction in texture definition. This presented an interesting possibility that the blurred regions actually contain class-specific information that is more important than was initially expected. Being formed from the underlying crystal structures of the original sample, the blurred regions whilst lacking in structural detail, still adhere to primitive pixel interactions that emulate those of the unaffected regions. See Chapter 5, section 5.3,

for information on the exploitation of these blurred regions for the extraction of potentially classifiable information.

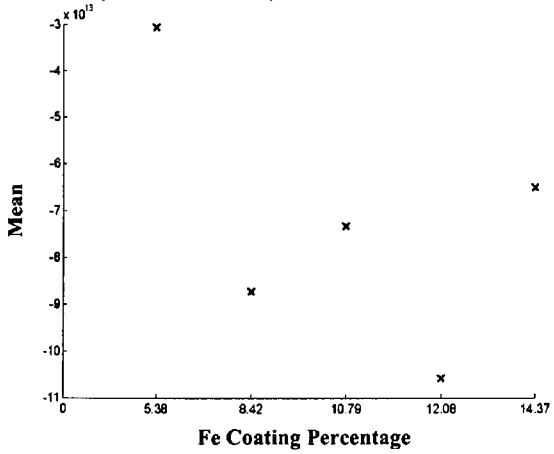
Figure 4.7: A selection of the results that enable classification of the galvanneal surface using a priori knowledge



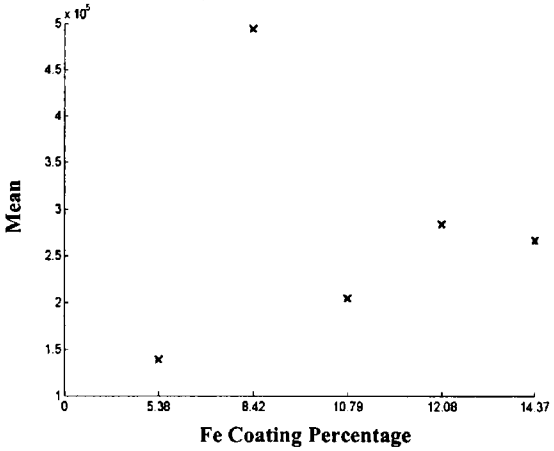
Mean (Correlation)



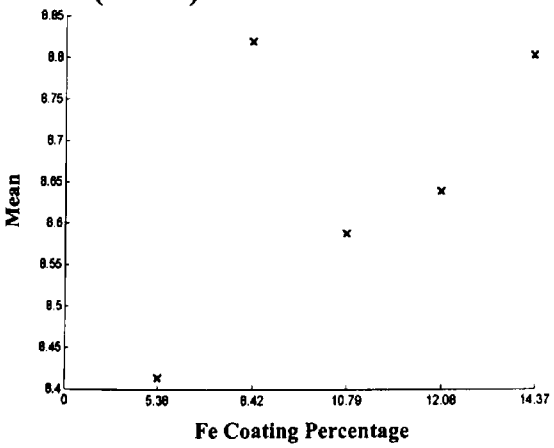
Mean (Inertia)



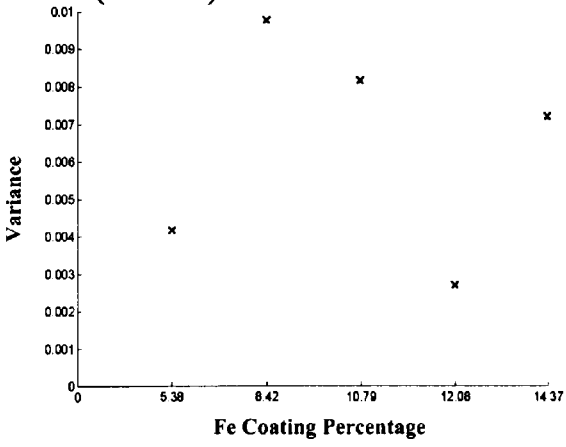
Mean (Shade)



Mean (EDMO)



Mean (Entropy)



Variance (Maximum Probability)

4.4.4 Conclusion

The grey level co-occurrence matrix is clearly not affected by the value of a single pixel to the same extent as the first-order histogram technique. Even so, viewing each two-dimensional feature in the spatial domain exposes a definite shift in the pixel mass for certain in-class images. This shift, either towards the lower indices for a darker image or towards the lower-right quadrant for an image with a higher average intensity, directly affects the GLCM measurements reliant on directional movement rather than the actual co-occurrence frequency count. Any minor differences within the in-class images are further exaggerated by employing the feature index as a weighting value for each respective frequency value. Nevertheless, the three aforementioned measurements (shade, prominence and correlation) which rely heavily on the value of the GLCM pointer generated some of the tightest clusters for the entire feature set. Whilst this implies that the variations in intensity and crystal formation for certain classes were not that great, their ability to classify the entire image set remains rather limited as there still existed a marked level of cross-cluster contamination. However, the analysis and combination of each cluster measurement with a priori knowledge, produced a promising set of results. At a classification rate of 100%, the feature has proved to be a capable performer that could potentially perform on-line at a more robust and reliable level than the probability density function.

4.5 Classifying Galvanneal Steel using Second Order Statistics – Without A Priori Knowledge

4.5.1 Genetic Algorithms

The genetic algorithm (GA) [4, 68, 69, 122, 123, 124] procedure is an iterative technique that attempts to search and tailor an initial population of individuals to a predefined problem. The GA utilises basic biological principles (crossover, mutate and selection [68]) to analyse the search-space and grade each individual according to a predetermined "fitness function" ("objective function" [125]). Individuals that produce strong fitness values are encouraged to "reproduce" and progress to the next generation, whilst the weaker ones are destined to be rejected and removed. Providing the fitness function accurately defines the desired result, the correct solution should be attained, although this is not always guaranteed.

4.5.2 Experiments and Discussion

After the promising a priori results generated by the ten GLCM measurements (section 4.4.2), it was necessary to examine their performance when using a trained classification system, and investigate their ability to perform as a potential on-line galvanneal classifier.

The five galvanneal steel classes were separated into two equal portions of 30 training and 30 test images. Ten measurements (section 4.4.2) were extracted from each GLCM generated with distance $d = 1$ and angle $\Theta = 0^{\circ}$. The GA procedure was used to generate a population of weighting values that would be used to draw in each of the ten GLCM descriptors to a measurement and class specific centroid.

Image class was defined on a nearest centroid basis rather than nearest neighbour due to the amount of overlap between clusters. This enabled the strict definition of each centroid as being the mean of each individual training class measurement spread. Such a value was deemed suitable enough to use in conjunction with the optimised genetic weights. Once defined, the weights would form the basis of a filter codebook that could, in theory, be used to classify any steel image of a recognised annealing type.

Experiments included:

- Varying the size of population to examine whether initial population size improves the efficiency of the GA search procedure [4, 126, 127].

- Varying the weighting values from 1.0 to 10.0, this offered an initial starting range that could be expanded or further restricted depending on their performance.

- Attempting various fitness functions to help refine successive element selection and to reduce the risk of premature convergence [128].

- Increasing the crossover and mutation parameters from an initial 0.6 and 0.03 [4, 126] respectively to help maintain a diverse population and to note the affect on consecutive generations.

- Implementation of population renewal and elitism schemes [126, 129], again to help maintain a varied population while the crossover and mutation parameters were being adjusted.

- Extension of the original population to include negative values, this would help to compensate for the wide variance either side of each in-class measurement centroid.

- Use of alternative centroid values to further increase between class distance and remove the reliance on a fixed training set.

- Development of "push-pull" filters to reduce in-class cluster variance, whilst expanding between class differences.

The GA weighting selection routine was implemented using the following four basic steps:

(i) Create an initial population of a predetermined number of randomly generated individuals.

(ii) Evaluate their performance (fitness) based on the critical objective function.

(iii) Place those elements that produce the greatest level of fitness into a mating subset.

(iv) Expand the search of the solution-space by applying the combination and mutation operators on the fittest individuals.

(v) Repeat steps two to four for either a fixed number of generations or until a predefined condition is met.

The initial population and indeed successive generations consisted exclusively of floating-point elements. Following the suggestions of [69, 74, 130, 131] the decision was taken not to waste CPU cycles, by eliminating the necessity for a phenotype to genotype encoding/decoding process [129].

There are two main limitations imposed by a finite alphabet with regards to a weighting generation scheme [132]. Firstly, it requires some form of prior knowledge regarding in-class measurement variance in order to ensure that the gene mapping of the upper and lower weighting limits, take into account the wide spread of values and differing scales between each GLCM descriptor. Secondly, by limiting the coding scheme to a select range of values, the accuracy of the weights is reduced when compared with the possibilities offered by the real valued equivalent. Even though the restricted alphabet may offer the advantage of eliminating the need to test any chromosomes within the search-space that would not form or help to produce a suitable weighting value, the advantage is weakened when it is considered that a bounded floating-point range could achieve similar results.

Whilst Goldberg [4] and Holland [68], Wright [133] and Antonisse [134], and Greenhalgh and Marshall [135] argue about the relative advantages and disadvantages of using binary digits, if the solution requires a floating-point value, then it would seem more sensible to base the GA population on the required type.

The decision to use a floating-point gene pool meant that an alternative form of crossover and mutation was required. Many crossover techniques exist for the binary gene, including 1/2 point, multi-point and uniform. For the mutate operation, bits are inverted based on comparisons between the value of a random variable and a predefined probability rate. A detailed description of each procedure and its purpose is presented in [69, 128, 129, 136, 137, 138]. With no bit-string to manipulate both operations must act directly on the numeric values. Using a random number generator and an equal probability of any method occurring at any one time, the following floating-point modification techniques were implemented [137]:

(I) Crossover

Offspring determined by:

(i) Calculating the arithmetic mean of the two parent genes.

(ii) Calculating their product and determining the square-root.

(iii) Subtracting the two values and adding it to the larger gene or subtracting it from the lower.

(II) Mutation

Offspring determined by:

(i) Adding/Subtracting a percentage value of the original gene.

(ii) Multiplying its value by a random percentage.

(iii) Replacing the original gene with a randomly generated alternative.

Applying certain combination and mutation schemes to a floating-point alphabet exposes the system to the possibility that the offspring may not be directly linked to its parents. This would produce a scheme lacking in direction and focus, and would offer no improvement over the basic random search. Therefore, each technique has been designed to mimic the mode of operation of the binary equivalent by attempting to retain the connection between the parent genes and their offspring. Nevertheless, the third mutation option of gene replacement introduces the possibility of analysing areas of the search-space that may not have been previously examined. Furthermore, the mutation system does not rely on the value of an individual bit, therefore, it has the potential to reach areas that could otherwise take a number of generations using a binary alphabet.

As stated in [4, 126, 127] the genetic algorithm's ability to perform an effective search is mainly determined by the size of the population. If the initial group size is too small then genetic diversity is lost and the end population tends to quickly settle around the same region. This problem, known as premature convergence [128], was found to remain even after alternative fitness functions were implemented. In fact, the search for the optimum weighting values relied solely on the combination of favourable mutations and the number of permitted generations. Conversely, maintaining a large population, whilst ensuring a varied selection of genes, reduced the speed of operation by forcing the examination of a greater percentage of (poor quality) genes.

The size of the population was set to 10 individually maintained groups of 100, with each cluster independently optimised based on the current GLCM measurement. Increasing the population size to 10 regions of 300 enabled the reduction of the number of generations required to produce the optimum set of weights, and reduced the need for both crossover and mutation, since the values required were already present in the initial population. Similar results were produced by the partial population renewal and elitist schemes. These included:

(i) Entire population replaced by its children on each generation.

(ii) Least fit $x\%$ of the population replaced by children (expanding on the elitism theory of retaining a single best gene from each parent generation [126]).

(iii) $x\%$ of the population replaced by random values.

Whilst they also allowed for a smaller population size and less frequent use of crossover and mutate, they offered little advantage over the standard techniques.

Returning the group size back to 100 and increasing the likelihood of both crossover and mutate functions occurring did not offer any tangible benefit. In fact, as the probability levels steadily increased, the level of gene fitness returned steadily decreased and showed no improvement through successive generations. This was because the good genes were effectively being destroyed and by the overuse of each operation useful individuals were prevented from fulfilling their potential. On the other hand with crossover and mutate probabilities set too low, the system ended up relying on the original population with poor

results, unless the number of genes and permissible generations were drastically increased to compensate.

Restricting the population size to values between 1.0 and 10.0 resulted in weights that were so limited in scope that they were unable to reduce those elements that either settled near the cluster boundaries, or possessed a value that was greater than the designated centroid. By allowing the initial population to contain both negative and a greater range of positive values this limitation was theoretically removed. This resulted in tighter clusters for each training class but only for very low sample counts (<10). When these weights were applied to the full set of test images it was noticed that they failed to correct the wide spread variance evident in the original GLCM measurements and in certain cases further aggravated the problem.

In a further endeavour to reduce the amount of between class crossover, the operation of the original GA fitness function, which was designed to reduce the distance between centroid and data point, was modified. The updated function attempted to triangulate the incoming training data measurement with the matching cluster centroid and the counterpart crossover data point from an opposing training class. It was instructed to maximise the distance between the centroid under focus and the crossover measurement, whilst minimising the range between the in-class element and its cluster prototype. The generation of weights which "push" opposing classes away whilst "pulling" the in-class measurements towards the key centroid, was introduced when the previous idea failed to produce clusters tight enough to prevent cross-cluster contamination. Two separate populations were maintained and individually graded based on their ability to either draw-in the matching data point, or repel the maverick measurement. Weighting pairs were

formed from the best gene from each population and collectively combined to form a "push-pull" library for each class. These libraries could then be applied to the incoming test measurements and attempt to mould the data points into clusters that possess clear, well-defined boundaries.

This new method also failed to correctly group the incoming data, even after various fitness functions and centroid values were employed. The results produced indicated that this new system was shifting the scale for the trained data and was drawing-in even more crossover measurements whilst pushing away boundary members. In operation the weighting-pair libraries were unable to distinguish between "good" and "bad" information.

4.5.3 Conclusion

With hindsight it is clear that in this case, the adjustment of any performance defining GA parameter was a purely academic exercise. The idea of using weights to reduce cluster size is fundamentally flawed when the majority of between class dimensions overlap, and the in-class variance can be considered too great. The net result is a system that is unable to yield a common set of weights that suit every in-class measurement. For this reason the GA technique was abandoned.

4.5.4 K-Nearest Neighbour

The K-nearest neighbour (Knn) classifier [26, 54, 119, 139] was employed as an alternative method of classification following the inability of the GA weighting optimisation routine to correctly identify the GLCM measurement points. Applying a

tried and tested classification technique shifts the emphasis from the actual method of classification, back to the analysis of the incoming measurements. It offers some indication of the suitability of the designated feature at representing the original image, and providing the feature is appropriate, the classifier provides some indication of image quality and homogeneity.

The Knn classifier is a commonly employed extension of the nearest neighbour algorithm [74]. Its approach to classification involves the assignment of an incoming data point, to a class that is best represented by the individuals surrounding its K-nearest neighbours. The classifier requires no pre-processing of the training clusters prior to element classification, and although techniques do exist that reduce the computational burden of Knn analysis [5], they were not practised here.

4.5.5 Experiments and Discussion

The 395 images were reduced to 300 and split into two equal groups of 150 for both training and testing. Classes B, D, F, H, J, represent the training group where each letter denotes a separate class based on the increasing amount of iron within the steel coating, and classes A, C, E, G, I, the corresponding set used for testing. The test data was measured against the training groups and the classification rates compared for the various cluster growth schemes. As with the GA classifier the ten GLCM measurements (section 4.4.2) were extracted from a matrix with distance $d = 1$ and angle $\Theta = 0^{\circ}$.

Experiments included:

- Varying the value of K to analyse cluster compactness.

- Various cluster assignment schemes.

    - Grow assigned cluster to analyse results produced by the standard Knn classifier.

    - Only grow cluster if correct class assigned, to determine classification effectiveness using a priori knowledge.

    - Do not grow cluster to assess training and test image similarities by retaining a fixed boundary based on the training class spread.

The value K can be considered as a measure of cluster compactness (density) and an estimation of cross-cluster contamination and boundary shape. These points were verified when K was steadily increased from 2 to 25 using the three cluster expansion routines. Rather than restricting K to only odd values (as a measure of avoiding the need for a tie-breaking decision scheme [140]), the casting vote was merely assigned to the nearest cluster. For the overall image classification vote using the ten GLCM descriptors (section 4.4.2), the sum of the difference between the point under focus and its surrounding K neighbours was calculated, and the final decision was based on the class that produced the lowest total over the ten dimensions.

Firstly, the standard Knn measurement was used to assign the incoming data to the cluster whose K elements can be considered as the individual's nearest neighbour (Table 4.1). As

the value of K was increased beyond 2, the level of classification steadily dropped right up to the 25 nearest neighbour limit. With the best rates of classification realised at the lower values of K, more specifically, when K equalled 2 or 3 the overall percentage was at 62%. A figure only produced because of the high individual classification levels for the extreme under-alloyed and optimum-alloyed images (5.38% Fe and 10.25% Fe respectively).

Table 4.1: The effect of varying K on the standard Knn classifier

K	% CLASSIFIED FOR EACH OF THE 5 ANNEALING GROUPS					OVERALL % CLASSIFIED
	5.38% Fe A	8.63% Fe C	10.25% Fe E	11.48% Fe G	13.96% Fe I	
2	100	63	93	46	10	62
3	100	53	93	53	10	62
5	100	53	93	10	0	51
7	100	53	93	3	0	50
9	100	53	50	3	6	42
11	100	53	56	3	6	44
13	43	50	86	0	0	36
15	43	46	100	3	0	38
20	50	50	96	3	0	40
25	50	50	93	0	0	38

The remaining groups performed poorly for two possible reasons. Firstly, because the Knn is unable to distinguish between a correct and incorrect classification without some form of prior knowledge, any incoming point is assigned to the nearest cluster whether or not it belongs to the same class. This leads to certain images expanding the boundary of a rival group, forcing one particular class to dominate any other within the same region. As more data points are added, the two rival clusters merge into a single mass, where the ability of the Knn to correctly classify any image is based solely on cluster density. This is reflected in the higher values of K where at 14-15, the extreme under-alloyed cluster (5.38% Fe) reaches its density limit and the classification rate falls from 100% to 43%.

The more densely packed optimum-alloyed class (10.25% Fe) remains stable throughout K and replaces the extreme under-alloyed as the dominant group.

Secondly, it was discovered that the Knn classifier is affected to a certain degree by the order in which the training images are analysed. In a similar vein to the previous problem, if the incoming data from one particular class favours an individual cluster, then the group has the initial advantage of being able to expand its borders unhindered (ignoring the other training clusters). Thus gaining in strength and further extending its reach by claiming incoming, bordering class data points as its own. This distorts the original classification problem, into one of attempting to accurately mimic the coating trend during a typical run. However, because the annealing process relies on environmental conditions that constantly vary, evidence of a distinct coating trend during a live galvanneal run remains non-existent.

The second method of only growing the cluster when the correct group was assigned (Table 4.2), was an exercise in examining how the Knn system would perform with the benefit of a priori knowledge. This method, as expected, returned the highest rate of classification at around 67% when K equalled 2.

As with the previous method, the number of images classified steadily decreased as K ventured beyond 2. This suggested that the problem of certain measurements crossing boundaries still existed even though the expansion of clusters using wrongly classified data was prevented. It also provided some indication that the classifier was not the only reason for the poor classification rates, but it was also the ten GLCM descriptors that were not providing enough distance between each annealing type cluster to allow the Knn

to cope with the in-class variances. Moreover, this method still allowed certain groups to dominate by allowing their borders to encroach upon those of their neighbours. Whilst restricting the growth of each cluster to only in-class elements reduced the density level around the borders and helped to increase the classification rates for the over-alloyed groups G and I, the level of crossover that still remained significantly dampened the value of possessing a priori knowledge.

Table 4.2: The effect of varying K on the a priori Knn classifier

K	% CLASSIFIED FOR EACH OF THE 5 ANNEALING GROUPS					OVERALL % CLASSIFIED
	5.38% Fe A	8.63% Fe C	10.25% Fe E	11.48% Fe G	13.96% Fe I	
2	100	70	93	63	10	67
3	100	60	90	60	13	64
5	100	46	93	40	6	57
7	100	50	70	33	23	55
9	100	50	66	10	20	49
11	100	53	56	33	10	50
13	43	70	56	13	23	41
15	43	76	50	10	16	39
20	50	63	60	13	13	40
25	50	70	50	3	3	35

Preventing any form of cluster growth allows for a direct comparison between the training and test data (Table 4.3) and eliminates the possibility of unfair group expansion. The results generated the second highest classification rate of 66% (K = 2) and verified that the original growth scheme was not always increasing the area of the correct cluster, but was allowing certain clusters to expand and dominate the other classes.

Table 4.3: The effect of varying K on the Knn classifier, using a fixed cluster size

K	% CLASSIFIED FOR EACH OF THE 5 ANNEALING GROUPS					OVERALL % CLASSIFIED
	5.38% Fe A	8.63% Fe C	10.25% Fe E	11.48% Fe G	13.96% Fe I	
2	100	56	83	73	20	66
3	100	53	73	76	23	65
5	100	40	66	76	30	62
7	100	46	80	76	23	65
9	100	50	80	56	23	62
11	53	50	86	53	26	54
13	43	50	73	56	30	50
15	43	53	73	60	30	52
20	50	56	66	43	40	51
25	50	53	66	46	46	52

Rather surprisingly this method of retaining a fixed border offered a very similar result to the second mode of a priori cluster assignment. This suggests that the values classified using the previous method actually increased cluster density rather than their size. The overall rates remain consistent throughout K, ranging from 50% to 66%. Even though this technique was unaffected by the order of the incoming test images, the shape and size of the original clusters had a dramatic effect on the individual classification rates for each annealing class. Reversing the roles of the training and test groups (Table 4.4) revealed a radical drop in performance for both the under-alloyed images of group B (using training subset A) and the optimum-alloyed group F (using training subset E). These results demonstrate an almost total reversal of the previous table (group I/J still performed poorly) and indicate the Knn's sensitivity to differences in size, density and shape. Factors that are difficult to control considering the fact that each individual pair of alternative clusters used for training and testing were both extracted from the same class.

Table 4.4: Reversing the roles of the training and test groups

K	% CLASSIFIED FOR EACH OF THE 5 ANNEALING GROUPS					OVERALL % CLASSIFIED
	5.38% Fe B	8.63% Fe D	10.25% Fe F	11.48% Fe H	13.96% Fe J	
2	3	73	43	76	30	45
3	3	80	56	73	56	54
5	20	86	53	73	46	56
7	16	86	53	76	46	56
9	16	80	50	70	46	52
11	16	80	56	76	56	57
13	16	83	63	80	60	60
15	16	80	63	73	60	58
20	16	83	60	70	50	56
25	16	86	66	63	50	56

It was decided not to hybridise the GA and Knn approaches [71, 74] because with ten dimensions/image, adjusting and re-adjusting the scale would serve only to shift the crossover boundaries rather than separate them. Of course, the number of dimensions could have been reduced [141] but no one measurement was sufficient enough to fully classify any one annealing group.

4.5.6 Conclusion

Based on the ten GLCM measurements, the Knn classification system failed to classify the galvanneal steel samples with any degree of confidence. Any classifier that is sensitive to both the order in which it encounters any of the five annealing types and the size, shape and density of the original training cluster, can not possibly offer a viable solution.

On a live on-line system, if the Knn classifier encounters a greater number of optimum-alloyed samples for example, there is an increased probability that an incoming neighbouring data point of a bordering class would be wrongly assigned the optimum

label. This would increasingly affect the Knn's classification performance as the optimum cluster's domination of the alternative groups begins to escalate. Whilst the option is available to prune and restrict the boundaries [5], doing so would not leave much scope for the variations evident within each class. Additionally, training the system after any alterations in the surrounding environment, would become a long-winded operation of determining the optimum prototype-cluster parameters as new image features present themselves. Any change in the capture environment would produce a collection of new cluster shapes and a new potential for class measurement crossover.

The Knn classifier and the ten GLCM descriptors are unable to produce separate clusters dense enough to enable accurate classification of the galvanneal samples and are of no practical use as an on-line combination.

4.6 Statistical Estimations and Inferences based on the GLCM Descriptors

4.6.1 Introduction

To further strengthen the conclusion drawn from the previous classification schemes, that the combination of blurred images and GLCM measurements are only of real use to a priori classification procedures, the data was subjected to a variety of statistical evaluations. This would provide a more detailed insight into the clustering arrangements of each of the ten dimensions, and give an indication of how these separate entities interact with one another. The Knn classifier allowed certain groups to dominate and

therefore only offered some indication that the measurement clusters were neither tight enough, nor isolated enough to be of any real use.

The statistical measurements allow the focus to be shifted to the GLCM descriptors and provide the means for an open examination of their performance rather than through the operation of a separate procedure.

4.6.2 Experiments and Discussion

The same 300 images as the Knn and GA classifiers were split into two groups, training (B, D, F, H, J) and testing (A, C, E, G, I). Ten GLCM measurements (section 4.4.2) were extracted from each grey level co-occurrence matrix based on distance $d = 1$ and angle $\Theta = 0^{\circ}$.

The following statistical measurements were applied to the entire data set:

- Number and percentage of images above and below the mean, to analyse cluster concentration around the central point (utilised as the GA cluster centroid for each GLCM measurement).

- Number and percentage of images above and below the mean that crossover from other classes, to evaluate each measurement's ability to uniquely categorise individual groups.

- Lower and upper limits of each GLCM measurement spread to assess the degree of crossover.

- Probability of assigning incoming data to the correct class, to assess the utility of each measurement for classification without prior knowledge.

- Validation of the Null Hypothesis that the GLCM measurements extracted from the low-quality galvanneal steel images, are only suitable for a priori classification. The following two distribution analysis techniques were used:

    - F distribution [142, 143] – 29^{o} of freedom, F .05. To compare the training and test cluster variance spreads, and to determine the level of in-class homogeneity.

    - T Test [142, 143] - Standard normal cumulative distribution function (CDF) – 95% confidence level used, to analyse and compare the mean cluster values for similarities.

The number of images above and below the mean varied for each class and for each GLCM measurement. The majority of clusters tended to be evenly spread at around 50-50, whilst the remainder demonstrated a 70-30 bias either way. The combination of this result with the fact that the upper and lower limits demonstrated a considerable amount of crossover served only to verify the problem of developing a common set of GA weights to draw each data point into a centrally positioned prototype.

Statistically analysing the variance and mean of each cluster was an exercise in confirming each measurement's ability to perform at an acceptable level with the assistance of a priori knowledge. The one-tailed F-distribution test was employed to compare two population variances, and to deduce whether the variance of one cluster was significantly greater or significantly less that the other. This would produce a similarity measure that could indicate the strength of the relationship between the training clusters and the incoming test data.

The findings summarised in Appendix 2, rather than providing any further insight into cluster interaction, served only to re-affirm the conclusions drawn from the original set of GLCM results. Even though the table summarising variance similarities indicate that a great deal of the clusters crossover (Appendix 2), it does not provide any indication of which class measurements span the same area. For certain groups the high frequency count of variance crossovers demonstrated throughout the range of annealing, represent a repeated intrusion of the original cluster space within the same GLCM measurement. The in-class variances of each cluster was considered to be well matched for specific dimensions, and certainly offered enough scope to classify the majority of the five annealing groups with a priori knowledge. A certain amount of difficulty arises from using a classifier that can not discriminate between "good" and "bad" measurements. Expecting it to deal with the wide range of measurement variations evident throughout the summary table is unrealistic. Similarly, comparing the cluster means using the standard normal CDF suggested a weaker link between the training and test data. It also offered further proof that using the cluster mean as a class centroid would not produce the most efficient set of classification results (pure speculation anyway, since the GA did not perform as expected).

The results derived from the analysis of the amount of measurements from class x that cross those of class y added further substance to the concluding remarks of the GA procedure. While the number of cases that crossed-over were maximised for the matching class training set for most cases, the differences realised were only minimal. This suggested that the cluster boundaries for all of the three classes of annealing were not well defined and tended to overlap, this verified the comments that were based on the results of varying K for the Knn procedure. As if to confirm this, the probability of assigning the incoming data to the correct class only managed a success rate between 25% and 35%. All the measurements produced a high probability of error irrespective of the annealing phase. The only possible exception was the extreme under-alloyed class, which produced respectable rates of classification for three out of the ten calculations. This supported the results obtained from the original GLCM work (section 4.4.3) where the same three descriptors, shade, prominence and correlation, produced individual clusters that were almost completely separable from the other groups. Using these calculations individually, it was possible to attain classification rates of between 70% and 80% for the extreme under-alloyed image set, but the overall results were weak.

4.6.3 Conclusion

As far as the a priori results are concerned, the statistical techniques indicated that in order to achieve the highest rate of a priori classification, the system should rely more on the variance measurements rather than the mean (even though in certain cases, both are required).

Whilst the statistical measurements failed to produce information that could have been useful in refining the quality of the original feature set, it still managed to verify the previous comments and conclusions. Irrespective of the fact that the classifiers failed to perform to the expected level, confirmation that the GLCM descriptors are clearly unable to separate each class without prior knowledge, has helped to define the cut-off point for the co-occurrence measurements, and has highlighted the need for an alternative feature set.

4.7 Interim Conclusion

Various feature extraction and statistical analysis techniques have been described and applied to the natural texture of the galvanneal steel surface. The successful application of these well-known methods by Haralick and Shanmugam [59], using the GLCM to classify reservoir sandstone, and Punch et al [71], using the GA and Knn combination to classify soil samples, was not repeated here. The low quality of the sample images, coupled with the relative sensitivity of the feature extraction methods to variations in steel crystal clarity, has produced a poor set of classification results. Attempting to improve upon the quality of the original galvanneal images through pre-processing, either using wavelet compression or software based filtering techniques, further exaggerated the problem of depth of field blur and significantly increased the level of perceivable noise. Similarly, the adjustment of any feature extraction parameter for any of the applied routines only improved the rates of classification to a limited degree.

Whilst the statistical measurements pinpointed problems with all three areas of the classification process, the source material, the feature representatives and the classifier,

the question that still remains is whether the images are at fault, or are the features used inappropriate for galvanneal steel classification?

4.8 Summary of Applied Techniques

(I) Wavelet Transform

(i) Impressive ability to compress images whilst maintaining basic crystal shape.

(ii) Did not offer any feature generation or classification benefits over the original data.

(II) First Order Statistical Pattern Recognition Techniques

(i) The probability density function is mathematically straightforward to calculate and contains useful information about grey-scale distribution.

(ii) Considered too sensitive to variations in lighting and environment to enable the distinction between annealing types.

(III) Grey Level Co-occurrence Matrix

(i) Not affected by variations in lighting to the same degree as the probability density function.

(ii) Well established technique that contains important textural information.

(iii) Promising set of a priori results.

(iv) A poor set of classification results, using two types of classifier.

Chapter 5

Galvanneal Steel Image Analysis using the Rank-Conditioned Morphological Transform and Novel Image Processing Techniques

5.1 Introduction

The failure of the previous methods to produce acceptable results for galvanneal steel classification has helped to determine the path for an alternative method of feature extraction and analysis. The rank-conditioned morphological transform offers a flexible mode of assessment and analysis that can determine crystal size and shape by using a basic geometric template.

The performance of an image feature is very much dependent on the extraction of information that accurately represents the data contained within the image region under focus. Whilst the GLCM extracted key formations from the galvanneal surface and provided possibly the most useful set of results yet, its non-a priori achievements remain poor. The two-dimensional feature lattice is considerably affected by the noise prevalent throughout the galvanneal steel image set. The matrix also demonstrates a shift in feature

shape and size in direct proportion to any variations in light intensity evident across the sample surface. This has led to the pursuit of an alternative set of image representatives that are less sensitive to the problematic blurred regions and offer the possibility of better representing the original sample.

Summary of applied techniques:

(i) Rank-conditioned morphological transform - using templates of varying size, shape and height, the crystal formations on the surface of the steel substrate was examined for any class-specific boundary shapes. At best, it managed to classify 71% of the galvanneal images but was unable to extract information from or around the blurred regions. Pre-processing techniques to improve probe access and increase the rate and accuracy of measurement extractions, were therefore developed.

(ii) Grey level co-occurrence filtration scheme - "opened up" blurred regions to enable an increased number of measurements whilst preserving basic crystal shape. In conjunction with spiral templates at varying height, the scheme improved the rate of classification to the overall maximum of 80%. However, the optimum height of the analysis template was difficult to determine. Consequently, methods for the determination of the height of a probe were investigated.

(iii) The design of template prototype intensity using the grey level co–occurrence matrix - the grey level co-occurrence matrix was used to analyse the pixel data encountered by each analysis structure using sample training images from each

galvanneal class, with the height of the template set according to the significance of the texture information extracted. The optimisation routine generated at best a classification result of 66%, an improvement over the 62% of the GLCM and Knn combination, but offering lower classification rates than the template intuitively designed by an expert user.

5.2 The Rank-Conditioned Morphological Transform

5.2.1 Introduction

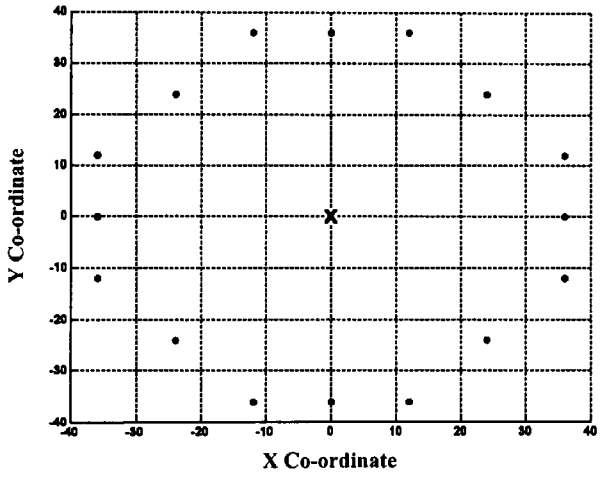
The Mathematical morphological approach can be considered as the alteration or assessment of image texture through the application of a basic probe structure. The structure can be considered as a group of connected resolution cells that collectively define a geometric shape with topological features (probe/template) [58].

Following on from [6, 76, 77, 78], the features used for steel classification are based on the rank condition measurement for the geometric shape to allow the point under test to remain part of the conditionally eroded result. The actual feature representative is formed from a set of histograms based upon the total number of probe templates applied to each image. The final examined feature is composed of the complete histogram set of the rank values required for each pixel-point in the sample to remain part of the conditionally eroded result for the given structuring template.

5.2.2 Rank-Conditioned Morphological Analysis

For an image $f(x,y)$, the process begins with the filter placed at the top surface of each point in the sample window $f(m,n) \in f(x,y)$, with respect to both the intensity level of the pixel at the probe locus (Figure 5.1) and the threshold level of each template prototype. Each pass of the image region generates a lattice that defines the level of containment over the entire structure at each locus position. In essence, this lattice represents the number of elements that did not contain information for each point in the sample window, and indicates the minimum rank required for the equivalent partial probe to be contained had a rank-conditioned (R-C) erosion been performed.

Figure 5.1: 2-D flat probe - Locus (origin) represented by centre point x

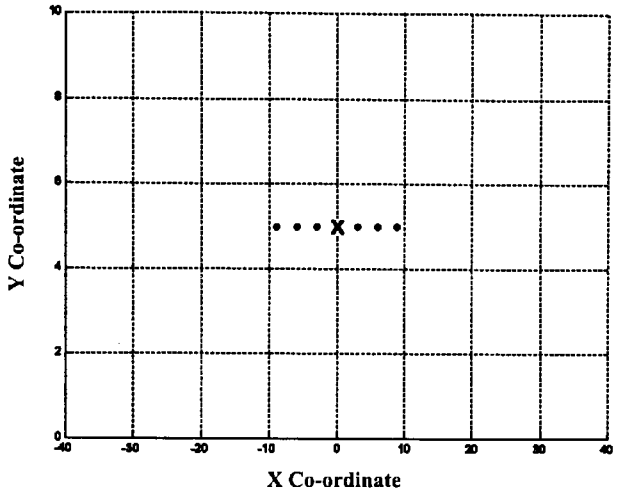


The image remains unaffected throughout the R-C operation since the process has been defined as a tool for texture analysis and comparison rather than image filtration and decomposition.

Texture classification using the rank-conditioned morphological transform begins with the setting of the intensity level of each individual element that forms the analysis

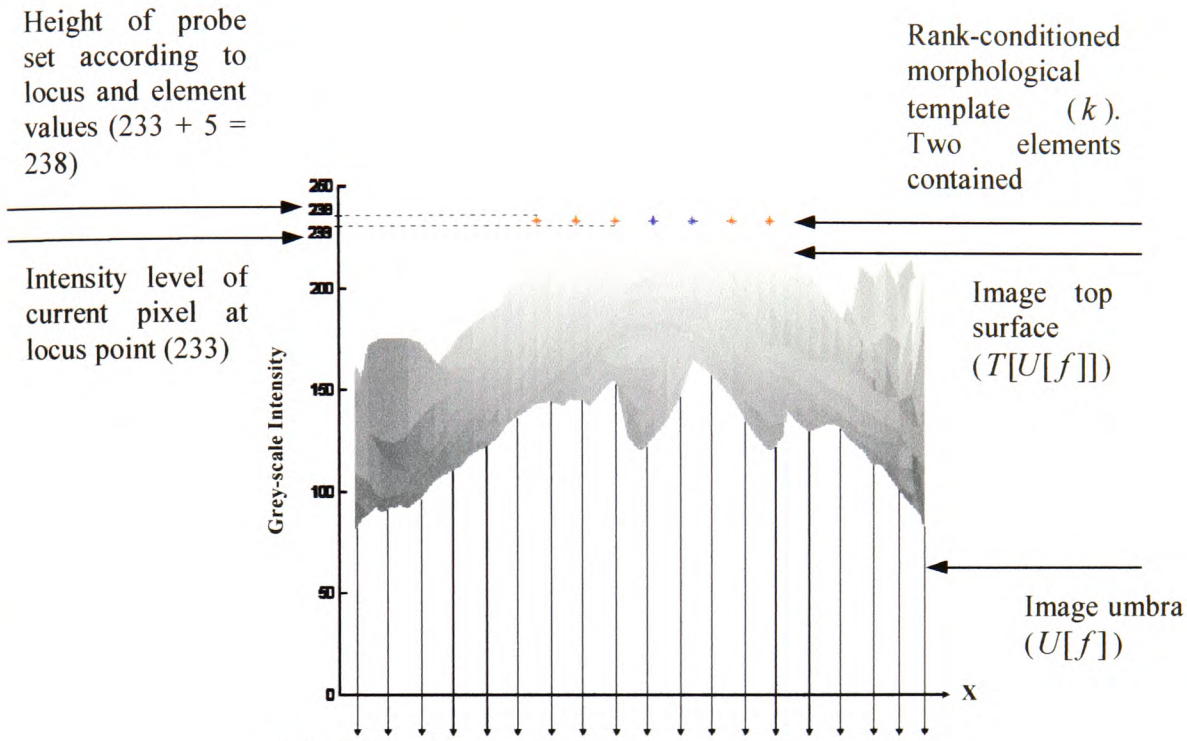
structure (Figure 5.2). Further work within this chapter emphasises the importance of setting the values to a suitable level in order to correctly classify the maximum amount of image data.

Figure 5.2: 2-D flat probe (side view) – Each element is set to an intensity level of 5. Locus (origin) represented by centre point x

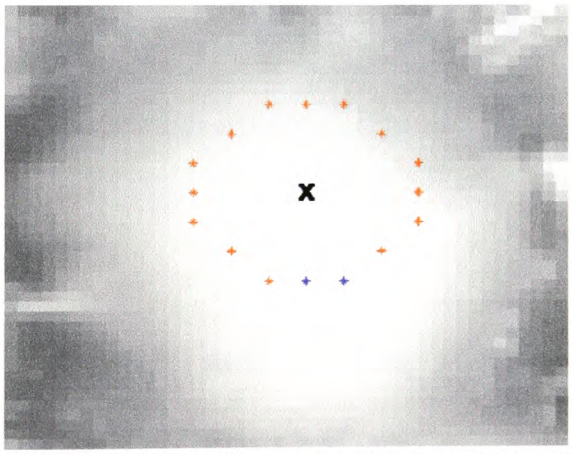


Next, the structure is placed on the top surface of the grey-scale image with its height set according to both the intensity level of each point within the probe and the grey-tone value at the locus position. A single stage R-C extraction is shown in Figures 5.3a and 5.3b. The example uses the template shown in Figure 5.2 (with the elements set to a threshold level of 5) and an image segment of 64^2 pixels extracted from an under-alloyed sample. Figure 5.3a illustrates that at the current locus position the pixel intensity is fixed at 233, by combining this number with the template threshold level of 5, the height of the probe is set to 238 (locus intensity + element intensity = height of analysis structure). The height of the structure varies as it traverses the image, constantly adjusting according to the grey-tone pixel value at the current locus position. For example, if the grey-tone at the next locus (pixel) point is fixed at 139, then the probe level would drop to 144.

Figure 5.3: Single stage R-C extraction. Key: Missed elements in red, contained elements in blue



(a) Side view

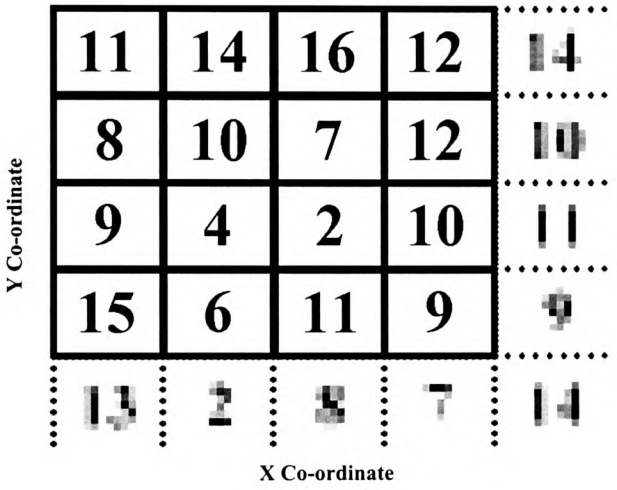


(b) Plan view. Locus (origin) represented by centre point x

Upon completion of each individual height adjustment, the pixels that reside at the position of each element is examined to assess the degree of containment. If the intensity value at any point is less than the threshold level of the probe, then the pixel is considered not to be contained within the structure. The actual feature used for classification (Figure

5.4) represents all the locus positions encountered by the template within the image region under analysis. At each location within the feature matrix (two-dimensional histogram) a value defines the number of elements not contained by the geometric structure. For example, at the first position in Figure 5.4 the template contained 5 out of the 16 elements, hence the number 11. Similarly, in the above example (Figure 5.3) 2 out of the 16 elements are contained at the current position within the sample image segment, therefore the equivalent point within the R-C histogram would hold the number 14.

Figure 5.4: Section extracted from a rank-conditioned morphological feature histogram



Discrimination is based on the variance values extracted from a secondary set of histograms representing the total frequency of containment for each template set. Using a minimum distance classifier each individual image within the test data collection is assigned to the class that is best represented by its nearest neighbour.

In mathematical terms, consider the top surface form of greyscale erosion for image f by probe k [144, 145],

$$f \Theta k(x) = T[U[f] \Theta U[k]]$$

where U represents the umbra and T the top surface [144, 145], (Figure 5.3a).

Physically interpreted as the minimum over the entire template,

$$f \Theta k(x) = \min_{z \in k} \{f(x+z) - k(z)\}$$

Now for any point (x_i) within the restricted sample window, set up to prevent boundary errors by taking into account the area covered by the applied filter,

$$(x_i) \in f(x) | (x_i + z) \in D_f$$

If,

$$([U[f(x_i + z)]] \geq [U[f(x_i) + k(z)]]) \forall z \in k(z)$$

then,

$$(f \Theta k)(x_i) = f(x_i)$$

This signifies that an individual image pixel at the probe element offset remains part of the conditionally eroded result, providing its intensity value is greater than or equal to the figure derived from the combination of locus intensity and the threshold value at the

current probe point (Figure 5.3). Relaxing this restriction to allow the pixel at $f(x_i)$ to belong to the eroded result for both partial and full containment over the entire structure requires that the assessment of template coverage becomes a measurement of the total number of elements contained for each template set (Figure 5.4). This is represented by the function @.

Giving,

$$(f @ k)(x) = R(x), j$$

where,

$$R(x) = \{r(x) : N, x \in f(x), (x + z \in D_f)|$$

$$r = \sum_{z \in k(z)} m, \quad m = 1, \text{ where } (f(x+z) \geq (k(z) + f(x));$$

$$m = 0, \text{ otherwise}\}$$

Where a $J(x)$ set generated for each $R(x)$ represents the level of missed containment for the complete set of resolution cells. The set is not used for the classification of texture and will not be discussed here (refer to [6, 76, 77, 78]).

The measurement feature then becomes a histogram H, based on the containment of the current probe when applied to the area analysed.

For each template,

$$H_i = \sum_{M,N} r(x,y) = i, i <= CARD(k(z))$$

where M,N represents the size of image region under analysis. The variance of each histogram is calculated and the linear distance between the generated feature descriptor defines the specific texture classes of the examined data set.

5.2.3 Experiments and Discussion

The structuring elements used consisted of a 2-D flat annular probe set (Appendix 3), sparse and filled (Figure 5.5). Both groups shared the same properties of increasing size and intensity, and offered the ability to assess 2-D texel size/shape. They were applied to the full set of 395 x500 magnified galvanneal images. In an attempt to remove the emphasis away from the blurred regions, the analysis window was fixed to almost the size of the original image, but restricted enough to prevent the introduction of unfavourable boundary conditions (Chapter 3, section 3.3, Figure 3.4).

Experiments included:

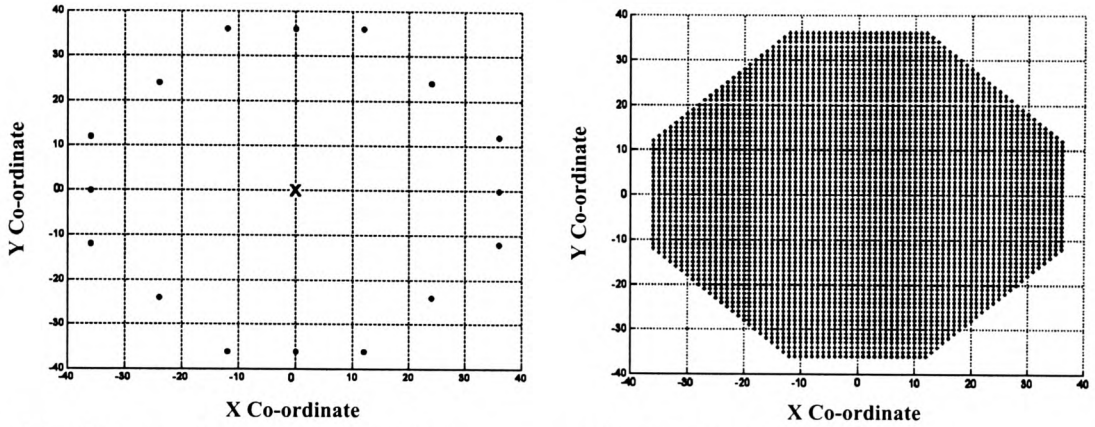
- Application of varying sized sparse annular templates (based on Figure 5.5a, outlined in Appendix 3) to assess the effect of probe size on pixel containment and class discrimination.

- The use of varying sized filled annular templates (based on Figure 5.5b, outlined in Appendix 3) to determine if an increased surface area would contain a greater

level of steel crystal data, and to compare its ability to classify the galvanneal samples against the sparse template equivalent.

Increasing the probe threshold values to determine whether the height of the probe has a greater affect on the rate of classification than its size.

Figure 5.5: 2-D flat sparse annular and filled probe samples



(a) 2-D probe (sparse annular) **(b) 2-D probe (filled)**

Table 5.1 illustrates the effects of adjusting both template size and intensity using the sparse 2-D annular probes and the x500 set of galvanneal images.

Adjusting the size of the probe, whilst maintaining the same threshold only produced significant differences when each probe was assigned a height of 0. With the classification rate dramatically decreasing as the size of probe increased. This implies that at the 0 level, the "small nut" probe was detecting the localised boundaries of the surface crystals and was making the best use of the regions unaffected by noise. Conversely, the "giant nut", by its very nature was unable to detect any crystal formations and was including a certain amount of blurred data at each point of containment. This meant that

the larger sized templates actually behaved in a similar manner at the 0 level, to the GLCM features that were generated at the higher displacement parameter values. However, where the co-occurrence feature was limited to the analysis of pixel-pair interactions at a basic two-dimensional level, the morphological template introduces a third dimension of height. This offers the possibility of extracting useful crystal information even though a large section of the probe may reside within the obscure areas. By increasing the threshold level of the probe beyond 0 (Figure 5.6), blurred regions are no longer considered as important as the surrounding structures. This enables the morphological feature to better distinguish between "good" and "bad" information than the original co-occurrence matrix. An increase in classification rate as the "giant nut" and "huge nut" templates are raised above the base level confirms the point. Additionally, the smaller probes "donut" and "small nut", only offer a slight increase over the lower equivalents, since they were already detecting the boundaries and were less affected by a combination of blurred region and crystal overlap.

Table 5.1: Summary of classification for various sized 2-D sparse annular probes, for the x500 image set

TEMPLATE INTENSITY	% CLASIFIED, X500 IMAGES
0	67
20	71
40	63
60	55
80	47
100	49
120	39
140	27
160	23
180	22
200	22

(a) "Small nut" template

TEMPLATE INTENSITY	% CLASIFIED, X500 IMAGES
0	62
20	71
40	65
60	59
80	58
100	53
120	45
140	30
160	24
180	23
200	22

(b) "Donut" template

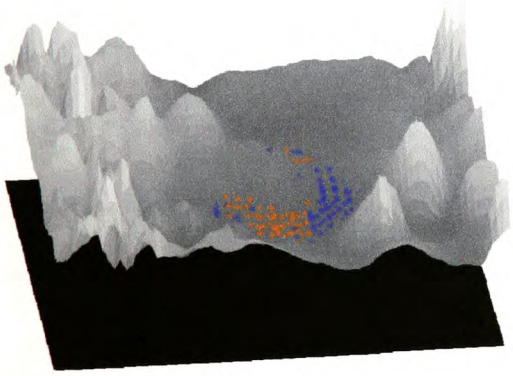
TEMPLATE INTENSITY	% CLASIFIED, X500 IMAGES
0	45
20	66
40	60
60	56
80	57
100	55
120	50
140	38
160	22
180	22
200	23

(c) "Huge nut" template

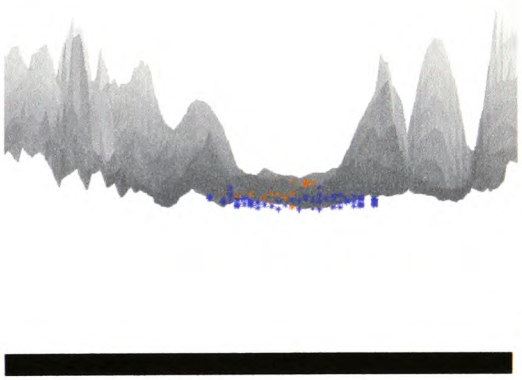
TEMPLATE INTENSITY	% CLASIFIED, X500 IMAGES
0	26
20	64
40	67
60	62
80	55
100	50
120	49
140	43
160	27
180	20
200	22

(d) "Giant nut" template

Figure 5.6: Varying the height of the probe to reduce the importance placed on the blurred regions. Key: Missed elements in red, contained elements in blue

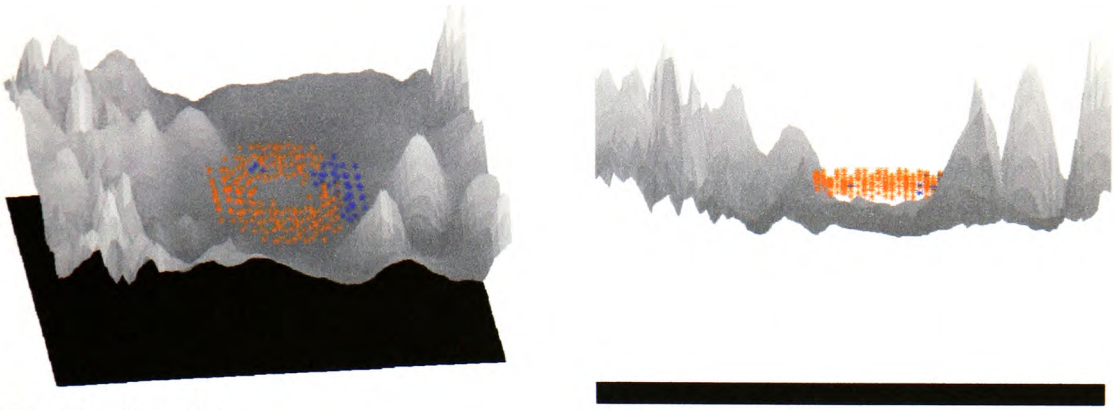


(a) Surface view



(b) Side view

At a height of 0, the template contains more of the crossover, blurred information, which undermines the benefit gained from containing the classifiable surface structures. (15 iterations, "Huge nut" template)



(c) Surface view **(d) Side view**

At a height of 20, the template now contains less of the crossover, blurred information, whilst still being able to extract crystal shape. (15 iterations, "Huge nut" template)

The "small nut" probe produced the best classification rate of 67% at the zero level with the "giant nut" yielding the worst at around 26%. This offered further proof that the smaller probe was encountering and containing more crystal formations than the larger template allowing a greater level of homogeneous data to be retained and correctly classified. Overall both the "small nut" and "donut" structures produced the best result of 71% at an intensity level of 20.

The application of filled 2-D probes to the galvanneal steel images, enabled the investigation into the use of an analysis structure that offers the ability to examine a greater surface area than the skeleton equivalent at any given time. Without exception, the performance of the filled probes was equal or slightly worse than that of the sparse original, irrespective of size or intensity. Also the substantial increase in the amount of points representing each template resulted in the morphological operation on average,

taking three times longer than the sparse version to measure each sample. The filled "small nut" probe matched the original with a classification rate of 71%, with the "donut" structure dropping to 69%. Whilst increasing the surface area analysed still enabled the crystal boundaries to be detected, the filled templates caused more of the blurred, crossover data to be contained. This offset the initial advantage gained from their ability to extract more detailed information on basic crystal shape, and resulted in the stunted performance for the fuller structure. The filled "small nut" probe was less affected because it was still able to completely contain the essential crystal patterns with a reduced probability of incorporating any of the blurred information into the measurement at the same time.

5.2.4 Conclusion

The previous set of experiments has defined the point at which the standard rank-conditioned morphological technique is unable to offer any further improvements in its current form. Nevertheless, considering the lack of overall detail within the galvanneal images, the classification limit of 71% represents a definite improvement over the GLCM and Knn combination of 62%.

Varying the design parameters of the analysis structure has offered some indication that the threshold level of the template remains a significant factor when attempting to isolate key features, and is something that was not fully explored in Rees' work [6]. It also demonstrates that any changes made to the template shape, in terms of size and surface area, does not produce the same level of variation in classification performance as the adjustment of the structure's threshold limit. A key point that will be further verified in Chapter 6, section 6.2.

5.3 Galvanneal Steel Image Filtering using the Grey Level Co-occurrence Matrix

5.3.1 Introduction

The majority of the steel images in their unprocessed form, whilst offering a certain amount of classifiable crystal information, fail to provide enough clear data to facilitate accurate feature extraction and measurement. In certain cases, up to 97% of the image has been subjected to a depth of field blur. This blurring not only affects the amount of fine-level detail that defines the steel crystal boundaries, but also hinders the performance of the rank-conditioned morphological transform.

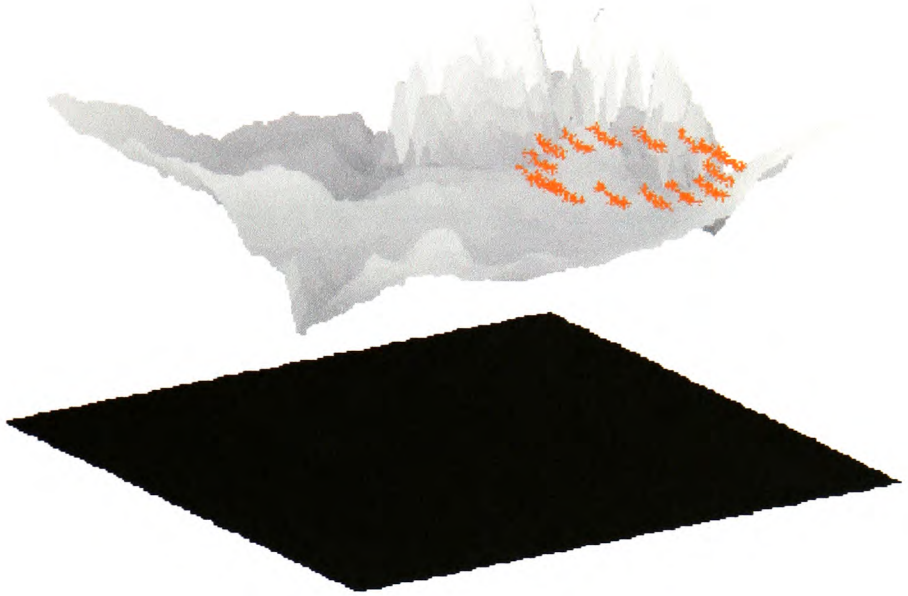
The R-C operation dictates that the probe level is set according to the grey-tone value of the pixel currently at the structure's locus (origin). If that particular value lies at such a point that the surrounding data is either below or around the same intensity level, then very little data would be contained (this also depends upon the threshold value of the original structuring element). The blurred regions of the sample images do offer little variation in grey-scale intensity, which results in the analysis template effectively "skimming" across the top surface (Figure 5.7a) and missing the underlying pixel arrangements. The analysis structure therefore, omits useful information that has the potential to increase the accuracy of the image feature representative. As shown in Figure 5.6, using a template of intensity 0 minimises the problem. However, at the surface level, the significant amount of noise prevalent within the blurred areas, serve only to aggravate the problem of image measurement crossover rather than improve the rate of classification. Nevertheless, by setting a fresh level at the base of the grey-scale range

within these "difficult" regions, the probe is able to contain far more of the previously ignored data (Figure 5.7b). Rather than adding additional noise by attempting to enhance or sharpen areas where no perceivable crystal shapes exist, the careful removal of pixel elements enhances the discriminatory use of these "cloudy" sections. By analysing the interaction between the new base level and the blurred surface, the measure of template containment becomes far more meaningful than the basic information extracted from the upper layer. This enhances the quality of the extracted measurements and generates a more accurate feature histogram than the unfiltered equivalent (see tables in section 5.3.3, for the performance differences between the filtered and unfiltered images).

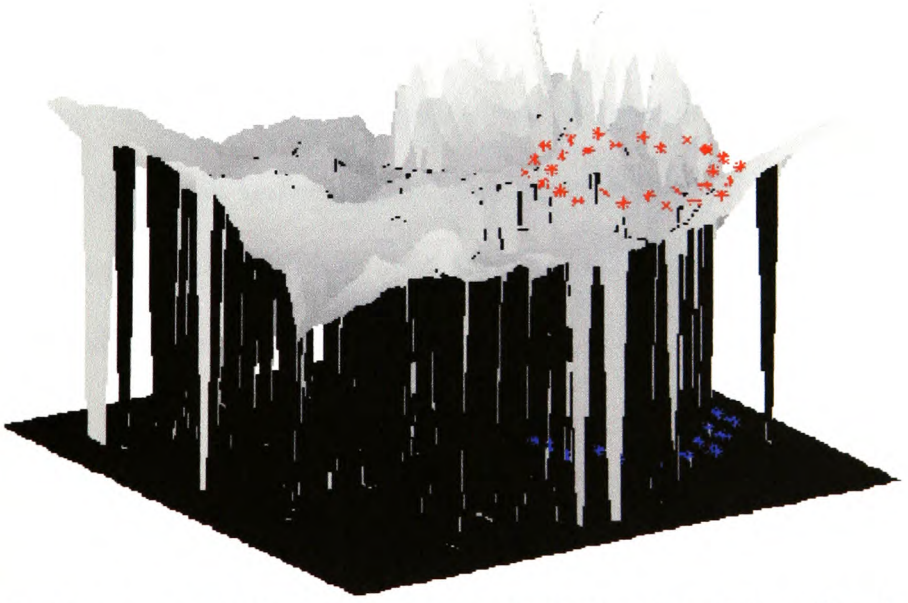
In addition to the choice of template shape and intensity, improvements in the rate of classification are determined by the decision to remove or retain a pixel in order to detect and enhance the subtle differences within these regions between each annealing group.

The establishment of a balance between the amount of blurred information to remove and the level of textural detail to retain, is the key to successfully "opening up" the image to an increased number of probe measurements, and hence an improved accuracy rate of the generated histogram. Over zealous filtering would not only damage class specific formations, but would also undermine the effectiveness of the morphological feature and its ability to accurately represent the original steel sample. Conversely, there would be little or no benefit over the standard non-filtered image, if the filtering operation were to be restricted to particular regions. Therefore, it is necessary to analyse the interaction between neighbouring pixels, and determine their importance based on their frequency of co-occurrence, rather than introduce an image refinement scheme that does not preserve the crystal boundaries and important texture arrangements.

Figure 5.7: **Comparing the level of containment for a 2-D sparse annular probe applied to the same region in non-filtered and filtered form Key: Missed elements in red, contained in blue**



(a) 2-D probe analysing a blurred, non-filtered region (5 iterations), unable to extract any key measurements



(b) 2-D probe analysing filtered equivalent region (5 iterations), now able to utilise the blurred information to exploit the between class differences within these areas and to improve the rate of classification

The idea of indiscriminately filtering an image and ignoring the neighbouring crystal relationships would fail to improve histogram accuracy for a number of reasons:

(i) Decision to filter an individual pixel is not based on image structure but on a non-connected random variable.

(ii) Different areas of the image affected on each separate application.

(iii) Each image in the class affected differently despite similarities in crystal formation and locality.

(iv) Crystal structures tend to be removed rather than "opened up".

(v) The process would not focus on the blurred regions.

Using the grey level co-occurrence matrix as a mechanism for decision based filtering, would address the above issues and enable the preservation of overall crystal definition. The result is a probe that has improved access to class specific material, including any areas that lack detail and focus.

Using one GLCM to represent each individual image did not take into account the local variations in crystal size and intensity. Transferring the results back to the original sample from the refined co-occurrence feature revealed a marked reduction in the amount of blurred information present but at a considerable cost to the useful textural data. Consequently, each image was dissected into equal sized regions to isolate the defective

material, and restrict filtration opening within the crystal boundaries. Window size analysis indicated that based on the GLCM filtration method employed, the smaller the region size, the weaker the impact. The filter window size, $r(m,n) \in f(x,y)$, was reduced from an initial $m\,x\,n = 150^2$, to $m\,x\,n = 50^2$ and finally to $m\,x\,n = 25^2$. The degree of modification to the acquired data steadily decreased at each reduction. Below this level, the filtering effect was insignificant, and failed to "open up" and reveal enough key features to show any improvements in the rate of classification.

Using a fixed window size of $m\,x\,n = 25^2$, the blurred data was targeted by exploiting its presence within the GLCM. It tended to be formed from a collection of frequently occurring pixels of similar intensity and at a level that was noticeably higher than that of the surrounding data. So, by removing the high frequency and therefore, high intensity GLCM elements, and transferring the results back to the original image, it was possible to open the cloudy regions to such an extent that they could be more accurately measured by the geometric structure. During any occasion whereby the GLCM was formed from a clear, detailed segment, then crystal damage would be minimised since the texels were typically constructed from low frequency components. Any filtering that did occur within these regions offered improved template access and produced a more precise measurement for the entire image.

5.3.2 GLCM Filtering in Operation

Taking an image $f(x,y)$, as function of grey-tone assignment, $f : x,y \rightarrow G$, where $G \in \{0...Ng-1\}$, where Ng represents grey-scale depth.

Giving GLCM,

$$P(i, j \in Ng^2 \mid d, \theta)$$

where, in this case $d = 1$ and $\theta = 0^o$, d is particularly important since only nearest neighbour pixels need to be examined.

For each window $r(m,n) \in f(x,y)$, where $m\,x\,n = 25^2$,

Gives,

$$P(i, j \in Ng^2 \mid 1, 0^o) \forall\, r(m,n) \in f(x,y)$$

Extracting the lower and upper threshold values, Tl and Tu respectively,

$$Tl = \min\{P(i, j \in Ng^2 \mid 1, 0^o)\}$$

$$Tu = \max\{P(i, j \in Ng^2 \mid 1, 0^o)\}$$

Calculating the central pivot,

$$Tp = (Tu - Tl)/2$$

Using Tl and Tp to produce the filtered GLCM Pf from P,

$$Pf(i,j) = \begin{cases} P(i,j), \text{where } ((P(i,j) < Tp) \wedge (P(i,j) >= Tl)) \\ 0, \text{otherwise} \end{cases}$$

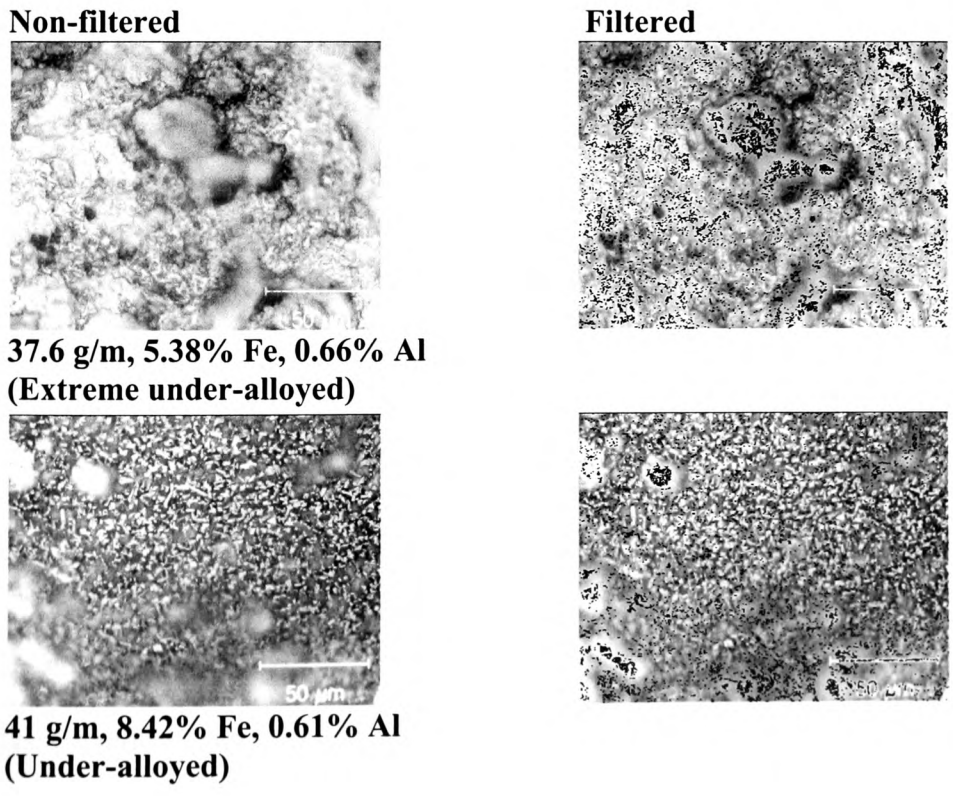
Transferring the filtered matrix Pf, back to the original image window. Where only the pixel-pair combinations that still exist within the filtered GLCM are retained at the same angle and displacement within the analysis region and the remaining couples are reset to zero.

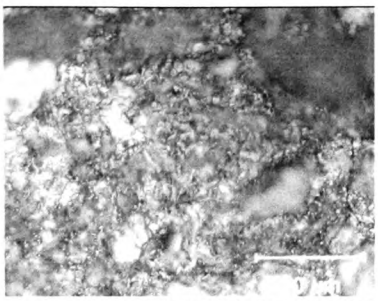
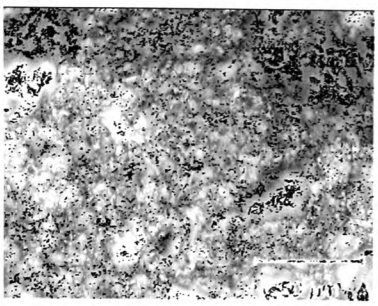
$$\forall r(m,n) \wedge r(m+d,n+\theta) \in f(x,y) =$$

$$\begin{cases} r(m,n) \wedge r(m+d,n+\theta), \text{where } ((Pf(i,j) > 0) \wedge (r(m,n) = i) \wedge \\ (r(m+d,n+\theta) = j)); \\ \\ 0, \text{where } ((r(m,n) = i) \wedge (r(m+d,n+\theta) = j) \wedge (Pf(i,j) = 0)) \end{cases}$$

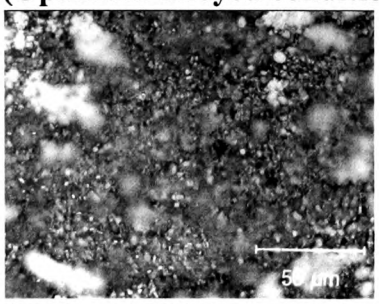
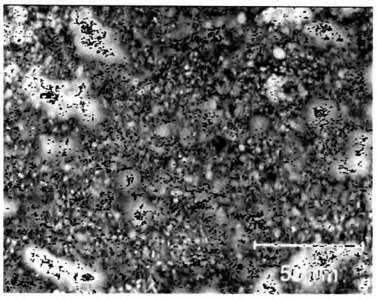
The net result is a galvanneal image that offers the measurement structure the potential to extract more useful information than the standard non-filtered equivalent whilst retaining the essential crystal structures (Figure 5.8).

Figure 5.8: Non-filtered and filtered galvanneal images

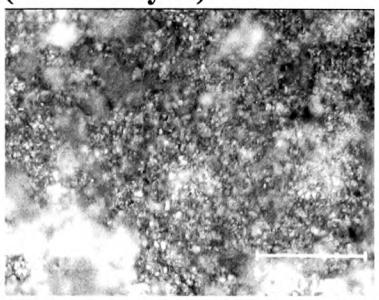
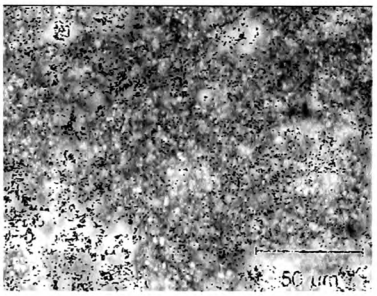


41.9 g/m, 10.79% Fe, 0.59% Al (Optimum-alloyed condition)

43 g/m, 12.08% Fe, 0.6% Al (Over-alloyed)

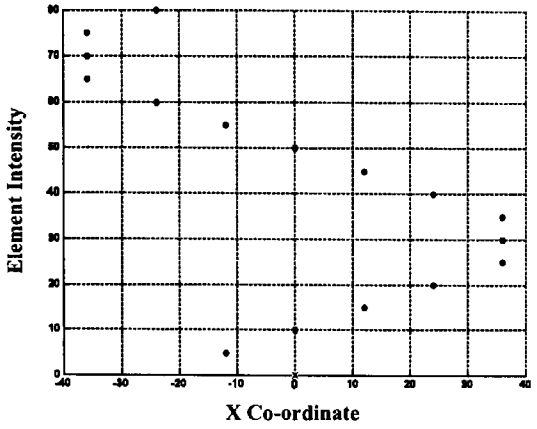
47.5 g/m, 14.37% Fe, 0.54% Al (Extreme over-alloyed)

5.3.3 Experiments and Discussion

The applied geometric structuring templates consisted of the previous sparse and filled 2-D sparse annular probe set, and a collection of 3-D single turn spirals (based on Figure 5.9, full set Appendix 3). Where the 2-D flat probes offer the ability to assess 2-D texel size and shape, the spiral templates enable a detailed examination of the 3-D texel profile which has been further enhanced by the GLCM filtration scheme. The three groups of templates were applied to the full set of 395 x500 magnified galvanneal images. Again, the analysis window was fixed to almost the size of the original image, but restricted

enough to prevent the introduction of unfavourable boundary conditions (Chapter 3, section 3.3, Figure 3.4).

Figure 5.9: 3-D Spiral template (side-view)



Experiments included:

- The use of varying sized sparse and filled annular templates (based on Figure 5.5, outlined in Appendix 3) to compare the differences in class discrimination between the filtered and non-filtered images.

- Application of 3-D single turn spiral templates of varying size and height, to introduce an extra dimension within the analysis structure that could better define the crystal structures from within the filtered regions.

- Increasing the probe threshold values to determine whether varying the height of the probe has a beneficial effect on the measure of containment for the GLCM filtered crystals.

When varying the 2-D probes intensity level the standard, non-filtered images offered worse results than the filtered equivalent (Table 5.2). Not just with the overall classification rate (although in the highest case this showed an improvement of more than 40%) but also the spread of above average classification values (50+%) throughout the intensity range.

Table 5.2: Summary of classification for various sized 2-D sparse probes, for both the standard and GLCM filtered x500 image sets

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	67	72
20	71	72
40	63	65
60	55	65
80	47	60
100	49	65
120	39	72
140	27	68
160	23	63
180	22	54
200	22	54

(a) "Small nut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	62	68
20	71	71
40	65	70
60	59	68
80	58	61
100	53	68
120	45	71
140	30	70
160	24	64
180	23	55
200	22	53

(b) "Donut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	45	48
20	66	67
40	60	67
60	56	65
80	57	67
100	55	73
120	50	74
140	38	68
160	22	63
180	22	58
200	23	52

(c) "Huge nut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	26	29
20	64	63
40	67	70
60	62	67
80	55	63
100	50	65
120	49	62
140	43	64
160	27	59
180	20	51
200	22	44

(d) "Giant nut" template

The results demonstrate the filter's ability to "open up" the standard steel image and enable an accurate measurement to be extracted using any shaped probe at varying intensity levels. In every case (with the exception of giant nut, intensity 20) the filtered image offered an improved rate of classification, with the gap between the standard and filtered rates widening as the template intensity level was increased. The "Huge nut" template at an intensity of 120 managed to classify 74% of the galvanneal images, the highest rate for the 2-D sparse annular probe set.

Analysing the larger surface area in any one instant using the filled annular probes produced similar results to the non-filtered images. The one exception, filled "huge nut" at a height of 120, improved upon the original by 4% giving an overall classification rate of 78% using the filtered images. By filtering the steel image samples, the probe is able to better utilise the increased surface area. Whilst previously it was able to extract useful crystal information, its progress was hindered by the increase in the amount of noise also contained. The filtering readjusts the level of the probe so that it is able to both contain the surrounding crystal data and exploit the subtle differences within the blurred regions when its locus resides within a filtered area. Nevertheless, the structure continued to require around three times longer to extract the necessary features, and still did not produce the highest rate of classification.

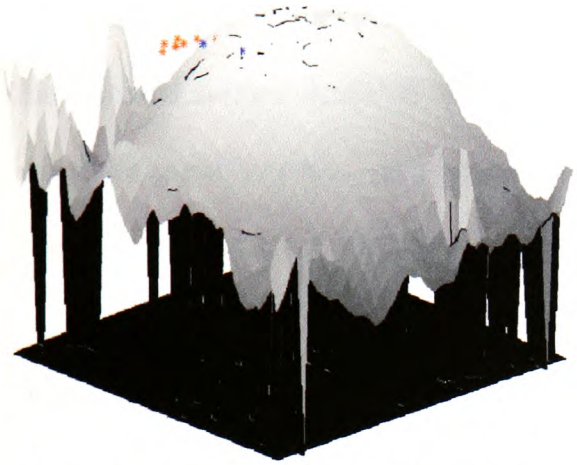
Changing the mode of analysis from the inspection of 2-D structure formations to the study of 3-D texel shape involved varying the intensity level of each individual template prototype. Using the 2-D sparse annular probes as a functional template, each prototype in a clockwise direction starting with the element at position, x = 12, y = -36 (Figure 5.5a), was assigned a progressively larger value. In this case the stepping value was a multiple

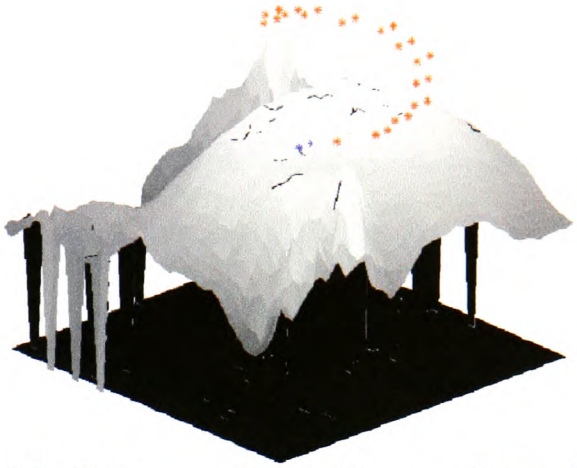
of 5, since it was better able to differentiate between the variations in intensity within the blurred regions. Where for example, a probe constructed from 16 elements would start at 0 and end at a height of 75. The next template in the same sized set would start at 5 and end at 80 and so on. The resultant spiral collection would present an additional measurement dimension by which surface features could be more accurately examined than the original 2-D templates.

Comparing the classification performance of sample 2-D and 3-D probes (Figure 5.10) when applied to a filtered segment, extracted from both extreme under-alloyed and extreme over-alloyed segments of steel, illustrates the benefit of using the single turn spiral.

Figure 5.10: Performance comparison of the 2-D sparse annular probe with the 3-D spiral template using GLCM filtered blurred regions extracted from two different classes of galvanneal image. Key: Missed elements in red, contained elements in blue

(a) Under-alloyed, "huge nut" probe **(b) Optimum-alloyed, "huge nut" probe**





(c) Under-alloyed, "huge nut" spiral **(d) Optimum-alloyed, "huge nut" spiral**

Figures 5.10a and 5.10b represent the "huge nut" template applied to both classes of steel. Analysing the level of extraction indicates that in both cases the structure is containing similar pixel information, despite fundamental topological differences it is unable to differentiate between the two groups. Conversely, the spiral probe (Figures 5.10c and 5.10d) produces two very different distributions. The variations in element threshold over the entire template, helps to promote the detection of the subtle differences in grey-tone presence within the blurred regions, and provides the means for the extraction of a more accurate measurement than the uniform equivalent.

As with the 2-D flat probes, the filtered images generated the highest rates of classification for the spiral templates (Table 5.3). The spirals were designed to improve the assessment of the GLCM filtered 3-D crystal profiles and to take advantage of the rising edge from the artificially created potholes. The rotating step formation of the geometric structure, resulted in classification rates that were significantly lower than the 2-D version when they were applied to the standard, non-filtered images. The noise-induced template "skimming" had a more profound effect on the 3-D structure because of

the increasing quality of each component's intensity level. This produced a probe that mimicked the operation of the flat templates over the blurred areas and contained less of the essential surface detail. Without the filtered regions, data was only being contained when the spiral was analysing valley-peak regions, where the surrounding peaks would rise above the structure's increasing height and enable at least partial containment.

Table 5.3: Summary of classification for various sized 3-D spiral probes, for both the standard and GLCM filtered x500 image sets

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	58	62
25-100	53	59
35-110	56	62
45-120	49	66
55-130	45	68
65-140	47	72
75-150	50	75
85-160	50	69
95-170	48	71
100-175	50	74
105-180	47	70
115-190	33	68
125-200	31	63

(a) "Small nut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	51	63
25-100	54	68
35-110	53	67
45-120	48	72
55-130	52	74
65-140	47	78
75-150	45	79
85-160	52	76
95-170	53	77
100-175	48	75
105-180	49	72
115-190	38	68
125-200	37	66

(b) "Donut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	54	63
25-100	54	67
35-110	52	69
45-120	52	71
55-130	55	75
65-140	55	78
75-150	55	77
85-160	54	77
95-170	48	79
100-175	46	77
105-180	53	76
115-190	48	74
125-200	44	69

(c) "Huge nut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	56	65
25-100	53	66
35-110	53	68
45-120	54	70
55-130	53	68
65-140	53	67
75-150	46	66
85-160	50	67
95-170	47	66
100-175	43	63
105-180	43	63
115-190	40	63
125-200	35	61

(d) "Giant nut" spiral template

The introduction of the GLCM filtering routine generated the necessary openings throughout the image to enable the spiral templates to extract accurate measurements from all regions not just valley to peak. The filtering-spiral combination generated results that surpassed those of the 2-D flat probes, not just in the overall rate of classification, but also in the spread of correct assignment values of 70% and above. This combination produced the most accurate level of image assignment with both the "donut" and "huge nut" templates producing a 79% accuracy rate.

Adjusting the size of probe whilst restricting the level of intensity produced similar results to the flat probe experiments. Even though it effected the overall result by a small margin, it was still possible to reach higher levels of separation by adjusting the intensity values to compensate for the new areas covered. For example the "donut" structure generates its highest rate of 79% at an intensity range of 75-150. Comparing this range with the larger templates indicate a classification rate of 77% and 66% for the "huge nut" and "giant nut" templates respectively. By increasing the height of the "huge nut" structure and lowering the "giant nut", their performance reaches their peak at 79% and 70% respectively. This offered evidence that for each size of probe, there exists within the image a set of valley to peak (high contrast) pixel interactions, both natural and artificially created by the filtering process, that help to distinguish between the five classes of galvanneal steel. A point that influenced the concept of creating a framework for the probe intensity optimisation scheme outlined in section 5.4.

Attempting to improve upon the rate of image classification beyond 79%, five spiral templates were applied to the steel images and the results combined. The probes were of the same fixed size but of varied height e.g. all small nut, intensity levels 0-75, 5-80, 10-

85, 15-90, 20-95 etc. For every image, the classification decision was based upon the overall vote taken for each individual nearest neighbour decision. The system returned an 80% classification rate when using the "donut" probes of intensity 70-145 up to and including 90-165. "Small nut", "huge nut" and "giant nut" combinations generated similar percentages but failed to improve upon the individual results.

5.3.4 Conclusion

Whilst the rank-conditioned morphological analysis technique demonstrated an ability to classify the natural crystal structures at a classification rate of 71%, it still ignored a great deal of classifiable surface information. In its standard, unfiltered form the blurred data offers no more than a collection of cloudy texture patterns that hinder the extraction of the clear crystal boundaries. Rather than attempting to ignore or completely remove these "damaged" regions which incidentally, would only leave around 30% of the image remaining in the majority of cases, accepting their presence and improving access to both them and the surrounding crystal formations has produced the best set of results yet.

The filtering technique was designed to retain basic boundary shape whilst "opening up" the image to an increased level of morphological measurements. This ability, coupled with the careful choice of spiral geometric structures produced a classification rate of 80%.

5.4 Morphological Template Optimisation for Texture Classification using the Grey Level Co-occurrence Matrix

5.4.1 Introduction

Strategies for the design of an optimal geometric template, in terms of shape and size for binary image restoration and noise removal are widely documented. Methodologies tend to revolve around mean-square estimations between the output of the filter and the desired signal [80, 81, 146, 147], or the application of a genetic algorithm search space routine [83, 84]. The techniques and results offer some useful insight into the possibility of an automatic probe generation routine that is better suited to detecting class specific texels than the non-optimised equivalent. Since morphological filters are utilised primarily as a tool for texture alteration, research in the optimisation of the structure for conditional morphological assessment and classification remains limited.

Classifying image sets that are comprised of grey-scale components introduces a problem not encountered in two-tone images. Where effective modification of the binary lattice is mainly determined by the size and shape of the applied probe (ignoring window size), the previous experiments have shown that grey-scale image feature extraction and analysis using rank-conditioned mathematical morphology relies primarily on the threshold level of each individual template prototype. Finer control can be exercised over the type and amount of grey-scale information contained, by varying the template prototype threshold level rather than the shape and size of the analysis structure. This has lead to an optimisation routine that attempts to define the optimal height of the constituent elements based on the variations in grey-tone pixel-pair interactions.

5.4.2 GLCM Analysis for the Definition and Extraction of the Optimum Morphological Template Parameters

Templates used for rank-conditioned (R-C) morphological measurements and classification, remain rigid, non-flexible structures whose parameters are selected usually through a process of trial and error. This leads to the extraction and analysis of information that is typically non-conducive to image classification. Table 5.4 illustrates the wide range of results generated for a fixed size "donut" template at various intensity levels, when applied to the galvanneal steel samples.

Table 5.4: Summary of classification results for the "donut" probe at various levels of intensity

TEMPLATE INTENSITY	% IMAGES CLASSIFIED
0	62
20	71
40	65
60	59
80	58
100	53
120	45
140	30
160	24
180	23
200	22

Through the direct visual inspection of each galvanneal class sample (Chapter 3, section 3.2, Figure 3.1c), it is not intuitively obvious that maximum separation would be attained by setting the probe threshold to 20. Therefore, multiple passes are required whilst varying the height of the analysis structure in order to attain the optimum result.

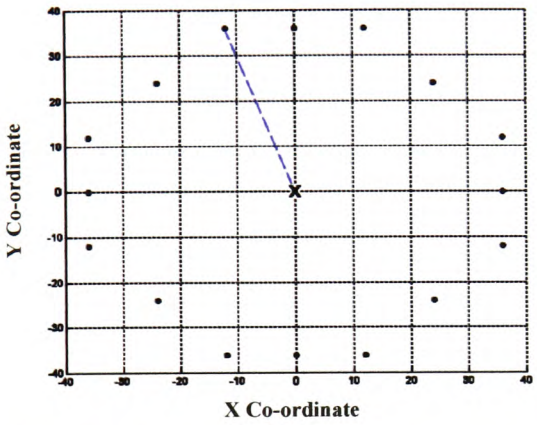
GLCM analysis attempts to remove the guesswork from template intensity selection and to always ensure that only class optimal information is analysed. The idea was influenced by the results obtained from varying the height of the spiral templates, where the intensity range of each element forming the structure can be adjusted to return the maximum rate of classification. The technique works on the assumption that for each group within the classification challenge, there exists intensity regions/texel formations that is either unique to the particular class, or contains more information at the specific levels than the alternative clusters. By extracting and utilising this information, an individual or a collection of probes of varying size can be optimised for a specific image set. The result is a number of templates for each class, each with their threshold level sensitive to the prototype patterns of the incoming training images. Varying the size of each probe within the template set provides an additional refined contrast measurement that assists the discriminatory capability of the individual analysis structure, and can be utilised to further detect differences between the crystal formations.

Define the origin of the structuring, or probe template as its locus, each individual structuring element being referenced to the locus both in position and level. The first stage of the threshold refinement scheme involves the treatment of each probe as a structured lattice of connected locus-element (l-e) pairs, Figure 5.11.

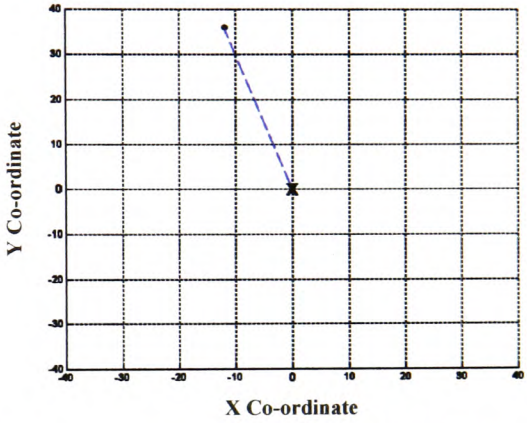
The R-C operation dictates that the probe level is set according to the grey-tone value at the pixel currently at the structure's locus. So, by generating a grey-level co-occurrence matrix using each l-e combination as displacement and angle parameters it is possible to examine potential probe containment for each individual element (Figures 5.12a and

5.12b). Once complete, this process continues on to the second stage of threshold extraction (Figure 5.12c).

Figure 5.11: Deconstructing a morphological template into locus-element pairs. Locus (origin) represented by centre point x

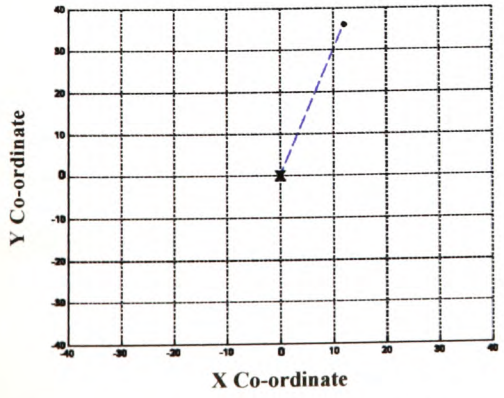


(a) The basic probe shape and size is selected. Each probe is separated into locus-element pairs, and used as GLCM displacement and angle parameters to evaluate pixel coverage for each individual element

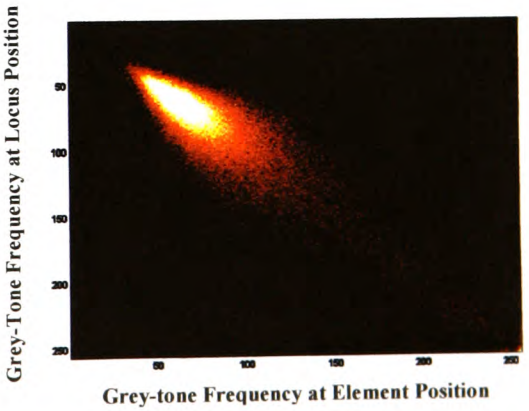


(b) Element 1 is connected to locus point x and is extracted to form the first l-e pair

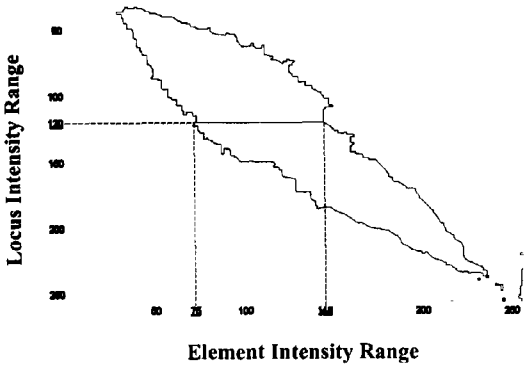
Figure 5.12: The four stages of GLCM analysis for probe parameter extraction



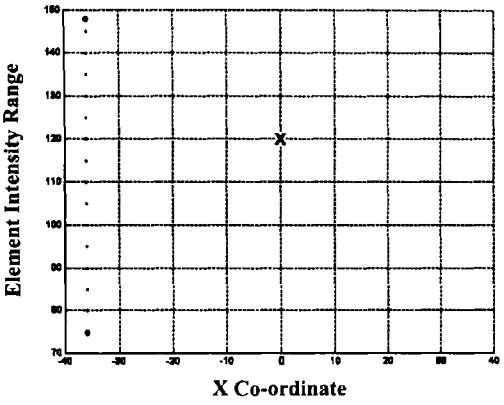
(a) Each probe is separated into constituent l-e pairs



(b) Each l-e combination generates an individual GLCM feature



(c) An element range is extracted for each locus value in turn (E.g. Locus 120 generates an element range of 75-148



(d) When locus point settles on an intensity value of 120, the variable element range is set to only accommodate pixels between 75 and 148 (for this example)

The element intensity values are extracted from the GLCM feature based on one of two selection criteria. Firstly, either the locus-element GLCM combination is present within the in-class training data at a frequency of occurrence that is higher than the alternative training groups, or secondly, the combination remains unique to the class under analysis. At this point, it is possible to formulate an optimum intensity level that would, in an ideal situation, only perform well on the corresponding class. Figure 5.12c provides an example of probe parameter extraction. As the probe locus encounters the grey-tone value of 120 for example, it is expected that for class x, that at the position of the element shown in Figure 5.12a the majority of pixel information would lie between the range of 75-148. On the condition that this sequence either does not exist or is not repeated at a higher frequency of occurrence in the alternative groups, then the intensity level for the first prototype would be set to -45.

The R-C measurement process indicates that an individual pixel at the probe element offset is contained, when its intensity value is greater than or equal to the figure derived

from the combination of locus intensity and the threshold value at the current probe point. Since each class has its own collection of morphological templates, it is important that each native structure be fully optimised for its individual group. Analysing the level of data containment based on template intensity indicated that the probe with the smallest threshold values always contained the most data (comparing fixed, same-sized templates). To counteract this undesirable property and to ensure that each probe set remains faithful to its own class, each template prototype is assigned a number of locus and element measurement decision switches (key-ranges). These switches are the pixel-pair ranges that are considered to contain most of the useful class-specific structural (texel) information based on the GLCM generated for each l-e combination. They are used to dictate whether or not a measurement should be extracted at the current l-e position within the image. Using the previous basic example for class x and the element from Figure 5.12a, a measurement would only be taken when the locus rested upon a grey-tone value of 120, and the pixel at the element position lied between 75 and 148 (Figure 5.12d). This would restrict analysis to only areas considered important to an individual class and to enable full containment within, rather than indiscriminately extracting measurements from the entire sample window. Any pixel combination that lies outside the designated range is appropriately penalised.

The key-range for each point in the probe becomes as varied as the available class specific information extracted from the GLCM of each l-e pair. Similarly the height of each probe automatically adjusts to accommodate all the available locus and element ranges and to ensure the maximum containment of classifiable information. The operation is identical to applying a collection of different structural templates at various threshold

levels to different intensity slices within the image. The approach restricts analysis to only areas of interest and to fully contain information unique only to its assigned cluster.

Once the elements have been optimised for each particular group, they are applied to the full set of test images. It was necessary to change the method of classification from the original rank-conditioned morphological probe histogram and nearest neighbour combination [76, 77] to the assessment of element containment through direct comparison. This works on the theory that if each prototype is optimised for a specific class, then it should contain more information than the elements tailored to the alternative images. Classification is then based on the discriminatory decision made by the optimised geometric structures rather than the current image's nearest neighbour. Measurements based on a template's constituent parts rather than using the structure as a whole, has enabled more accurate comparisons to be made between each group, allowed each locus-element pair to utilise their own key-ranges and help to further improve intensity assignment precision. If a reasonable degree of control over illumination can be obtained, the inclusion of this information should enhance the classification capability of the analysis process.

5.4.3 Selection of the Optimal Template Intensity Values

Specifying the process for a single template k and defining an individual image $f(x, y)$, as a function of grey-tone assignment, $f : x, y \rightarrow G$, where $G \in \{0...Ng-1\}$, where Ng represents grey-scale depth.

Giving GLCM,

$$P(i, j \in Ng^2 \,|d, \theta)$$

Now consider each training image as a set of points within a window restricted to prevent boundary errors with respect to the elements that form the applied filter,

$$(x_i) \in f(x) \,|\, (x_i + z) \in D_f$$

Using the current pointer for the image index $f(x_i)$ as representing the probe locus, in conjunction with each prototype point at z, to give an indication of pixel intensity coverage for each locus-element pair gives GLCM Pc,

$$Pc(i, j \in Ng^2 \,|dist_{f(x_i) \rightarrow f(x_i + z)}, \theta_{f(x_i) \rightarrow f(x_i + z)})$$

For each template there exists a set of GLCMs Ps, that defines the grey-tone combinations encountered by the entire structure on a point-by-point basis, where n represents the number of elements that form the probe (Figures 5.12a and 5.12b),

$$Ps = \{Pc_0, ..., Pc_{n-1}\}$$

The option is now available to filter the generated feature set or to continue without any modification. The filtering process enables the reduction in the amount of potential key-range values. This shifts the emphasis to in-class similarities rather than between-class

differences. By restricting the amount of measurements taken, the classifier is forced to examine only a limited selection of key pixel combinations that could be deemed as being suitable representatives for the entire class. On the other hand, by not reducing the range of measurement decision switches available, key features become less and less important as the set of probes optimised for one particular class start containing unwanted data from the alternative groups. This leads to the problem of the classifier expecting the incoming test image to exhibit the same properties of a specific class over the entire feature matrix, whilst displaying only limited similarities between itself and the alternative clusters.

Filtering option 1: Removing elements with frequency of occurrence $< x\%$ of the maximum occurring pixel-pair combination over the entire GLCM

Extracting the upper frequency value for the most commonly occurring pixel-pairs from each GLCM $Pc \in Ps$, constructed from the locus-element parameters,

$$Tu = \max\{Pc(i,j)\}$$

Using Tu, as the main reference value and removing elements $< x\%$ to prune the outer edges of the main diagonal gives the filtered matrix Pf,

$$Pf(i,j) = \begin{cases} 0, where\ Pc(i,j) < (Tu.x\%) \\ Pc(i,j), otherwise \end{cases}$$

Filtering option 2: Removing elements with frequency of occurrence $< x\%$ of the maximum occurring pixel-pair combination on a line-by-line basis

Extracting the maximum frequency value Tp, one row P_i at a time, from the current GLCM Pc, where $P_i = (P_{i0}, ..., P_{ij}) \in Pc$,

$$Tp = \max\{pi\}$$

Reducing its value by half to retain the larger frequency elements and to soften the effect of the filtering,

$$Tp = (Tp / 2)$$

Removing elements on a row-by-row basis that are $< x\%$ of the current Tp value, to give the filtered matrix Pf.

$$Pf(i,j) = \begin{cases} 0, where\ P_i(j) < (Tp.x\%) \\ P_i(j), otherwise \end{cases}$$

The result of applying either filtering method is a collection of image specific GLCM features Psf, that represents the filtered equivalent of the Ps set,

$$Psf = \{Pf_0, ..., Pf_{n-1}\}$$

The next step, irrespective of whether or not the filtering routines are applied, is to isolate those elements that contain potentially useful textural information. These provide the locus-element measurement decision switches to restrict template analysis to only key regions. For basic labelling purposes, it is assumed that only filtered matrices Pf are used.

Continuing the specific one image and one template case. Each row index i (representing the grey-scale values encountered at the locus position for the current l-e pair) is individually assessed to determine whether a measurement would be worth extracting given the amount of information registered for the current locus row. Giving a set of locus $L(r)$ values for each point in the probe,

$$L(r) = i, where(\sum_{pi \in Pf} P_i > 0)$$

where P_i refers to each row of the filtered lattice, $P_i = (P_{i0}, ..., P_{ij}) \in Pf$.

To further restrict probe access to specific textural information and to help formulate a dynamic element intensity value that adjusts to contain the maximum amount useable image data, each prototype $L(r)$ set possesses an accompanying lower El and upper Eu element pixel-range value. These help to specify the point of analysis within an image by first comparing the grey-tone at the current locus position with those stored in $L(r)$, and secondly comparing the pixel intensity level at the current element position with those stored in the corresponding element range $E(r)$. The element range is extracted from the same filtered GLCM set Pcf using $L(r)$ as the row index (Figure 5.12c),

$$E(r) = \{El, Eu\} \forall r \in L(r)$$

where El represents the first column j, in the GLCM to register any pixel-pair co-occurrence at any retained locus positions $L(r)$,

$$El = \min\{j\}, where(P_i(j) > 0) \forall_i \in L(r)$$

and Eu signifies the last column j, in the GLCM to register any grey-tone pixel-pair co-occurrence at any stored locus positions $L(r)$,

$$Eu = \max\{j\}, where(P_i(j+1) = 0) \forall_i \in L(r)$$

giving each probe element at point (z) its own set of intensity values based on the difference between each locus key value $L(r)$ in turn and its accompanying lower element intensity El,

$$k(z)(r) = \{L(r) - El\} \forall (z \in k) \wedge \forall (r \in L(r))$$

The resultant template prototypes exhibit an intensity threshold value that varies according to the current locus pixel (Figure 5.12d). Each point only extracts a measurement when both the locus and element pixels lie within the predetermined key-range library (example shown in Table 5.5).

Table 5.5: Template library format example

X	Y	LOCUS VALUE	ELEMENT LOWER RANGE	ELEMENT UPPER RANGE	TEMPLATE THRESHOLD
-30	10	175	180	190	5
		180	130	160	-50
		...	...	...	...
		...	...	...	...
		190	167	200	-23
-30	0	177	189	205	12
		...	...	...	...
		...	...	...	...
		185	178	215	-7

The mode of classification now changes from the combination of the original histogram and nearest neighbour analysis routine, to the comparison of element containment for each class optimised probe.

Expanding the image set from one image, to two separate classes $C = \{f_0(x_i), f_1(x_i)\}$, and $D = \{f_0(x_i), f_1(x_i)\}$, both containing two images, and the probe sets $Tc = \{k(z)\}$ and $Td = \{k(z)\}$ both having a single non-optimised template, places a new demand on the locus selection routine.

Again, analysing each row index i of every GLCM within the current Psf set for both classes C and D, where $P_i = (P_{i0}, ..., P_{ij}) \in Pf \; \forall \, C$ refers to each row of the filtered lattice for class C, and Pk the equivalent for class D, generates a locus range $L(r)$ for every element within the probe sets Tc and Td,

$$L(r) = i, where((\sum P_i > 0) \wedge (\sum P_i > ((\sum P_k) + (\sum P_k.x\%))))$$

Rather than simply registering the existence of a locus value by analysing the row and assigning i based on its summation value, additional comparisons must be made. A locus value is only retained as a key-range locus element if it either exists uniquely within the focus class[1] or if its frequency of occurrence is greater than the opposing groups by $x\%$.

The image classifier now takes the form of the following,

If the pixel intensity value at the current locus and element positions $f(x_i)$ and $f(x_i + z)$ respectively, are contained within the key range library for any probe,

$$([U[f(x_i)]] \in L(r) \wedge [U[f(x_i + z)]] \in E(r))$$

then a vote is taken for the entire probe set of both classes based on the difference between the pixel intensity at the element position and the summation of the optimised prototype intensity value $k(z)$ and the locus point $f(x_i)$,

First the summation S over every element $z \in k$,

$$S(c) = \sum abs(f(x_i + z) - (f(x_i) + k(z)))$$

Now, the vote is cast for the class that has missed the least amount of grey-scale data over the entire template,

---

[1] Since each locus value is treated equally, $\sum P_i > ((\sum P_k) + (P_k.x\%))$, implies that $\sum P_k = 0$

$$R(c) = R(c) + 1, where\, S(c) = \min\{S(c)\}$$

where $c \in \{0...Nc-1\}$, where Nc represents the total number of classes. If the locus grey-tone lies outside any of the predefined locus ranges then the minimum summation for the current locus-element class is penalised by the addition of the specific value for $\max\{G\}$, i.e. the maximum permitted grey-tone value,

giving class I for the probe that missed the least amount of information and therefore a greater number of votes for $R(c)$,

$$I = c, where\, R(c) = \max\{R(c)\}\forall c$$

The original rank-conditioned method simply defined overall probe coverage by incrementing the appropriate position within the generated feature histogram. Since each probe has now been optimised for its own individual class this mode of analysis is unsuitable. The R set now monitors the level of containment for each particular point in the probe and registers the class that produces the best results, with the incoming image being assigned to the I cluster whose members within the probe set, generated the highest number of votes.

5.4.4 Experiments and Discussion

Twenty-five (five/group) 2-D flat annular probes of varying size (based on Figure 5.1/5.5a), each with sixteen elements (with the exception of the "small nut" probe, which was formed from twelve points), were split into locus-element pairs and applied to an

initial twenty-five training images, five from each class. All l-e template combinations were used to generate separate GLCMs for key-range and intensity extraction. These switches would restrict analysis to only class specific combinations of locus and element intensities, with each range possessing a threshold value optimised to contain the maximum amount of characteristic data for each individual group. The optimised structures were then applied to the full set of 395 galvanneal steel samples that had been captured at a magnification of x500.

Experiments included:

- The use of alternative training images to determine whether the ranges used to decide if a measurement should be extracted (referred to as, key-ranges or measurement decision switches) remains constant and does not affect the rate of classification.

- Variation of the number of training images for both filtered and unfiltered GLCM matrices to examine growth changes in the measurement decision switches, i.e. do the number of measurements extracted increase or decrease according to the number of training images used.

- Adjusting the level of filtration for both filtering schemes to restrict the number of measurements extracted and to shift the focus to in-class similarities by only using only a portion of the available locus-element ranges from the GLCM feature matrix (Figure 5.12c).

- Application of probe optimisation technique to the GLCM filtered galvanneal images introduced in section 5.3 to test its generic classification capabilities.

Using sparse annular probes of varying size and intensity (0-200) in conjunction with the standard rank-conditioned morphological measurement, produced a classification rate of up to 71% (Table 5.6). However, a problem arises when the optimum threshold setting for any of the applied templates is unknown. In the case of the first probe "giant nut", the results indicate that the structure has the potential to classify between 20-67% of the incoming galvanneal images. Without repeated applications whilst adjusting the height there is no indication that 67% would be the optimum result. Similarly, with a large difference of 47% between the best and worst classification rates, it is essential that the probe level be accurately adjusted according to each individual case. Admittedly, the lowest rates of classification were observed towards the latter portion of the available grey-scale range, where the probability of containing significant discriminatory textural features is greatly reduced. Nevertheless, even with the possibility of extracting key class information from a choice collection of training samples, the selection of a template threshold level remains a non-trivial task.

Table 5.6: Classifying x500 samples using various sized sparse annular 2-D non-optimised templates (non-filtered images)

APPLIED TEMPLATE	% CLASSIFIED (LOWEST)	% CLASSIFIED (HIGHEST)
"Giant nut"	20	67
"Huge nut"	22	66
"Donut"	23	71
"Small nut"	22	71

The class optimised probes overcome the need for repeated applications by extracting all the necessary a priori information from a fixed number of training images. However, experiments in training the system using different samples from within each class suggest that variations in intensity produce a problem that is similar to that of the original non-optimised method as shown in Table 5.4. Table 5.7 illustrates the results generated from using five alternative groups of images and indicates the problem of attempting to produce a collection of templates individually tailored to each class. Examining the probes generated within each group has highlighted the variations in intensity within the training images that were apparent during the first and second order feature extraction techniques. This has produced differing l-e ranges for each in-class training set and has had a detrimental effect on the their usefulness as a measurement decision switch for the entire set of test images.

Table 5.7: Using alternative sets of five training images to illustrate the direct effect on the rate of classification (unfiltered feature matrix)

SET	% CLASSIFIED
1	48
2	50
3	58
4	61
5	64

With any system that relies on both training and test data, where the in-class members do not conform to a strict set of rules regarding quality and intensity range, it is important that the training samples be truly representative of the class. Otherwise the template prototype threshold values calculated from the extracted ranges of the chosen training set, may not be at the optimum level for the entire group. As far as the x500 galvanneal steel images are concerned the best representatives may not actually be the samples least

affected by the depth of field blur, but the ones that exhibit the most common set of properties whether they are "damaged" or "undamaged". The aim is not to produce the illusive "gold" template, but a structure that is best optimised to contain as much class specific information as possible.

Increasing the number of training images from 1/class to 20/class produced a marked growth in the number of key-range values, as the number of in-class image intensity variations expanded (Table 5.8).

Table 5.8: Varying number of training images (unfiltered feature matrix)

NUMBER OF TRAINING IMAGES	% CLASSIFIED
1	17
3	32
5	64
10	46
20	58

However, beyond the peak classification rate at 5 images/annealing group, the classifier failed to improve upon the 64%. This reduction relates directly to the amount of measurement decision switches available for each probe. As the number of key-range values increase, so does the likelihood of extracting cross-class measurements. This limits the effectiveness of the genuinely useful l-e keys as they become insignificant amongst the crossover members. Conversely, too few training images and the templates become too reliant on a small texture region within their set and depend upon the incoming data greatly favouring the properties of one particular class over the other alternatives. This problem worsens as the texture region under analysis reduces in size.

The filtering of the l-e feature matrix prior to range extraction was an exercise in evaluating the effect of range reduction. Tables 5.9 and 5.10 outline the results for the two previously defined filtering options for the x500 galvanneal image set.

Table 5.9: Filtered GLCM using filtering option 1, data set 5 training images/class

X%	% CLASSIFIED
10	53
20	50
30	52
40	51
50	50
60	37
70	56
80	17
90	17

Removing elements with frequency of occurrence $< x\%$ of the maximum occurring pixel-pair combination over the entire GLCM

Table 5.10: Filtered GLCM using filtering option 2, data set 5 training images/class

X%	% CLASSIFIED
10	66
20	62
30	64
40	65
50	66
60	65
70	66
80	66
90	62

Removing elements with frequency of occurrence $< x\%$ of the maximum occurring pixel-pair combination on a line-by-line basis

Comparisons against the unfiltered GLCM equivalent templates (Table 5.7), demonstrates that the shift in focus to in-class similarities by reducing the effective element key-range, has produced an improved level of classification. Filtering the GLCM based on the overall maximum value produces a far harsher level of element removal. Not only does it decrease the perceivable range of element values, but also greatly reduces the available measurement key switches. This produces a potentially dangerous situation whereby the reduction of the size of the analysis window, i.e. the number of measurements extracted from each image class, forces the process to rely upon a collection of much smaller regions where any detrimental effects caused by noise are greatly amplified. The severity of the filtering is controlled through the adjustment of the variable percentage x. As the value of x decreases (for both schemes) the number of GLCM elements increases. Throughout the percentage range the filtering severity of the first option produced only a maximum classification rate of 56%. Conversely, the line-by-line filtering method consistently generated results of 65% and above throughout the percentage range. Whilst this figure does not improve upon the original rank-conditioned non-optimised method, which gave 71%, the number of repeated applications necessary was significantly reduced. Nevertheless, the lowest result returned when using the alternative groups of training images (Table 5.7) bettered the lowest percentages of the original. Giving an indication that the optimised template is not as greatly affected by variations in image intensity as the non-optimised equivalent.

Increasing the number of test images with the filtered GLCMs (Tables 5.11 and 5.12) produced a similar set of results to the unfiltered version (Table 5.8), with the maximum rate of classification fixed at 66% as opposed to 64%. The lighter, shallower filtering of the second option (Table 5.12), offered a slight reduction in the width of the element

measurement range (x-axis, Figures 5.12b and 5.12c), and therefore increased the focus on the in-class similarities. This subtle form of range restriction secured the line-by-line filtering option as the best performing intensity selection routine, and produced the highest classification rate of 66% for the entire collection of threshold optimisation experiments. It offered a suitable balance between the wide element ranges extracted from the unfiltered GLCM and the widespread, severe filtering of the first filtering option, which unrealistically expected a high level of in-class uniformity.

Table 5.11: Varying number of training images using filtering option 1

NUMBER OF TRAINING IMAGES	% CLASSIFIED
1	17
3	31
5	56
10	52
20	53

Removing elements with frequency of occurrence $<x\%$ of the maximum occurring pixel-pair combination over the entire GLCM

Table 5.12: Varying number of training images using filtering option 2

NUMBER OF TRAINING IMAGES	% CLASSIFIED
1	17
3	44
5	66
10	54
20	60

Removing elements with frequency of occurrence $<x\%$ of the maximum occurring pixel-pair combination on a line-by-line basis

Applying the template optimisation routines to the GLCM filtered images that were first introduced in section 5.3, was an examination into how the system would perform on a pre-processed image set. If the system has potential as a generic classifier then the results should mimic those of the original, untouched samples. Adjusting any of the optimisation

parameters, including template size, filtering technique and number of training images had no effect whatsoever on the final conclusion. That is, the classification results remained almost exactly the same for both sample sets with differences of only 1-2% either way. This indicated that the template optimisation technique was utilising the image grey-tones to optimise the probe threshold levels to the best of its ability.

5.4.5 Conclusion

The perceived difficulty of predetermining a suitable template threshold level that is universally capable of extracting the prime texel formations, is understandable considering the variety of both pixel patterns and template parameters available. The experiments of section 5.2 have already indicated that by not specifying the optimal threshold level, classification performance can vary to a great degree. Similarly, the spiral probe experiments of section 5.3 illustrate that by incorrectly setting the spiral height its effectiveness is considerably reduced, irrespective of the filtering status of the galvanneal steel images. Clearly, the illusive, "magical template" that is "all things to all men" does not exist. Therefore, the creation of a process that is able to generate the most efficient set of threshold values for any shape or size of template and for any classification task, is highly desirable but has yet to be fulfilled.

Based on the previous set of results, the template intensity optimisation framework outlined is at present unable to compete with an equivalent template intuitively designed by an expert. With the key-range selection routine, the GLCM filtration scheme and the updated classifier having yet been fully explored and optimised. Nevertheless, it is still considered that combining 3-D shape, intensity and size into a single morphological measurement would produce the best classification results, for any form of image texture.

5.5 Interim Conclusion

The rank-conditioned morphological transform as first introduced by Rees and Jones [76] and further defined in [6, 77], has been applied to the galvanneal image set in an attempt to overcome the lack of crystal detail that greatly affected the techniques of Chapter 4. At 71%, the R-C operation produced a noticeable improvement in the rate of classification over the previous best of 62% generated by the GLCM and Knn combination. Visual inspection of the morphological feature extraction process (Figure 5.6) demonstrated the amount of information that was ignored by the geometric template, whether it was intentional (Figure 5.6c and 5.6d) or as a direct result of the interference caused by unnatural noise (Figure 5.7a). Whilst the generated histogram is accepted as being characteristic of the image surface, the blurred regions produced a feature that could not be wholly relied upon as being a true representative of both image and class. In cases such as this, where it is not possible to re-sample the steel substrate to minimise the amount of image noise, best use must be made of the available data. The information neglected by the R-C and sparse annular template combination, tended to be the regions most affected by noise. These areas would either be completely overlooked or interfere with the accuracy of measurement extraction where a blurred region would overlap a clear crystal boundary. Consequently, the cloudy regions were targeted for filtering using a scheme based on the interaction between pixel-pairs as defined by the grey level co-occurrence matrix. Using the same sparse annular templates and the newly filtered galvanneal steel images, the classification performance increased to 74% and by tailoring the probes to the filtered images, the rate increased by a further 6%.

The experiments to optimise the intensity level of the geometric structures in section 5.4, failed to improve upon the 80% of the non-optimised single turn spirals. With a best of

only 66% the routine offers scope for improvement. However, an intensity optimisation framework exists for further expansion and adjustment that in its final form could have the potential to classify any form of natural texture. Suggested improvements are outlined in Chapter 7, section 7.3.

The question of whether the image or feature is at fault still remains unanswered. Whilst, the majority of morphological experiments returned a rate of classification that surpassed those of the previous techniques, an insufficient amount of alternative image and texture types have been analysed to apportion the blame. Therefore, experiments performed in the next chapter go some way to rectify the lack of evidence and help to reach a final conclusion.

5.6 Summary of Applied Techniques

(I) Rank-Conditioned Morphological Transform

(i) Not as affected by the problem of depth of field blur as the previous set of techniques.

(ii) A flexible method of data analysis and image classification that could theoretically classify any form of image texture.

(iii) Ignores a great deal of potentially useful, classifiable image data.

(II) GLCM Image Filtering

(i) A filtering operation that maintains basic crystal boundary shape, whilst increasing the level of template access.

(ii) Combined with the single turn spiral probes, the operation produced the highest rate of classification at 80%.

(iii) A generic filtering routine that improves template access for other types of natural texture (Chapter 6).

(iv) Reduces the probability of producing a poor classification result, by maintaining probe access to the texel formations at all levels of template thresholding.

(III) Template Intensity Optimisation using GLCM Analysis

(i) Has the potential to optimise any shape or size of template for any type of image within the classification challenge.

(ii) Is not as significantly affected by the blurred regions as the original rank-conditioned morphological transform or the alternative techniques.

(iii) Unable to perform to the same level as a fixed template designed by an expert.

Chapter 6

Final Analysis and Discussion

6.1 Introduction

The feature extraction and classification algorithms of Chapters 4 and 5 have been implemented and applied to the full set of galvanneal steel images. The techniques of Chapter 4 are well defined and have been utilised on a number of other problems involving natural texture formations in many other literary examples (Chapter 2). Their ability to perform as a procedure for feature extraction and analysis is well proven, as a consequence they provide a useful comparative base for novel ideas. The methods are further evaluated and compared against those of the previous chapter. The on-going theme of technique versus image quality for the problem of galvanneal steel assessment sees the inclusion of the x200 galvanneal images, in addition to further experiments and procedure comparisons, to eliminate the sources of difficulty in classification.

Summary of experiments and chapter outline:

(i) Grey level co-occurrence matrix and K-nearest neighbour approaches - Applied to alternative image textures to evaluate their performance on non-blurred texel

arrangements. At a classification rate of 93%, the results for the x200 image set suggest that the technique is sensitive to the depth of field blur present within the x500 set, since the affected x500 images were classified to only 62%. However, the peak classification rate of 31% for the undamaged Brodatz subset may suggest otherwise.

(ii) Rank-conditioned morphological transform - Applied to the x200 and Brodatz images to assess how the obscure regions influence the final classification result. Classification percentages of 100% and 76% respectively indicate that the blurred areas may have a more profound effect than previously theorised. However, variations in image and class count could offer an alternative reason for the increased rate of classification over the x500 samples.

(iii) Conclusion and discussion - Summarises the results of the performed experiments and draws together the main points of the image quality versus technique debate.

6.2 Alternative Images and Comparison of Results

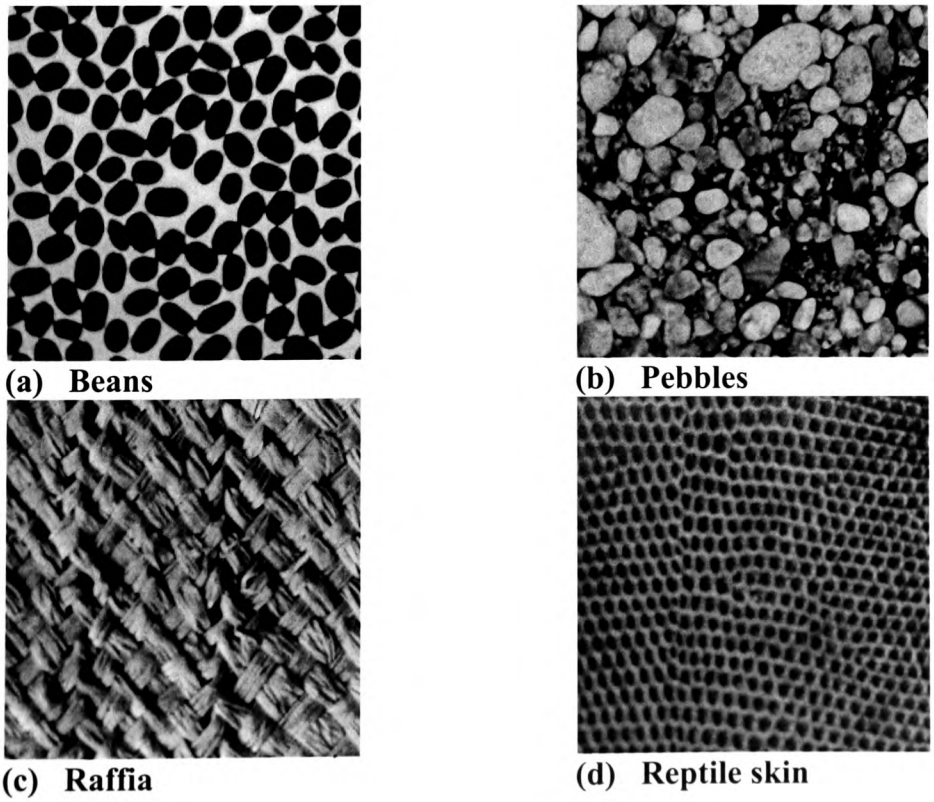
6.2.1 Introduction

Using various examples of test data either in the form of steel images or the natural texture samples drawn from the Brodatz set (see Appendix 4 for examined subset) [148], and analysing algorithmic performance, highlights the strengths and weaknesses of the applied techniques and suggests a definite progress path for future study (defined in Chapter 7, section 7.3).

The x200 galvanneal steel image set (examples from each class shown in Appendix 5) provide only a limited view of the fine crystal detail but lack the depth of field blur that was omnipresent throughout the x500 samples. At the lower level of magnification, the Leco 2001 image capturing system is well within the limits of the optical lens, avoiding the intermittent blurring caused by weaknesses in the lens at its performance boundary. These images not only offer comparative evidence of image utility for the detection of the steel annealing phase, but also highlight the unavoidable trade off between crystal size and image clarity.

The Brodatz samples (Figure 6.1) provide a stable foundation on which to compare the established GLCM and Knn approaches, with the novel morphological routines of the previous chapter.

Figure 6.1: Samples taken from the Brodatz set



(a) Beans **(b) Pebbles**

(c) Raffia **(d) Reptile skin**

6.2.2 Experiments and Discussion

The x200 steel images (Appendix 5) consist of 120 samples divided into 3 groups of 40. As with the x500 galvanneal set, the 3 groups are divided according to the annealing phase of the steel substrate, under-alloyed, over-alloyed and optimum alloyed. The x200 and x500 sample sets share the same image dimensions of 499 x 386 pixels and are restricted to the same 8-bit grey-scale range.

The Brodatz subset used for this project comprises of 18 classes of natural texture formations, e.g. beans, reptile skin, pebbles etc. (Appendix 4). Each image is 256^2 pixels in size, with each pixel 8-bit in depth. With only 1 sample/class, 9 smaller 128^2 pixel regions have been extracted from the original sample to increase the number of test images available. However, a certain degree of texture overlap occurs within each of the 9 images for each individual class. This is unavoidable considering the size of the source sample and the number of sub-images extracted. The final Brodatz test data collection is formed from 18 classes, 162 images of 128^2 pixels at a depth of 8-bits/pixel.

Experiments included:

- Application and comparison of the performance of the GLCM and Knn approach to texture extraction and classification, with the morphological (templates of varying size, shape and height), and the GLCM filtering and optimisation techniques.

- Utilisation of the above methods on the following image sets:

- x200 galvanneal steel images, filtered and unfiltered. The objective is to determine the ability of the GLCM filtering routine to improve template access for images unaffected by blurred regions. The best method of classifying the x200 images is evaluated, and the classification performance compared to the x500 results.

- Brodatz samples, GLCM filtered and unfiltered, to judge the capability of the selected feature extraction routines as applied to textures other than steel, and with a data set that possesses an increased number of texture classes.

- 3 individual sets of 120 samples, drawn from the original x500 galvanneal collection, and separated into their respective annealing classes of extreme over-alloyed (around 14.37% Fe), extreme under-alloyed (around 5.38% Fe) and optimum alloyed condition (around 10.79% Fe). Used to assess the level of damage caused by the regional blurring by comparing the results with the x200 collection of galvanneal images.

Table 6.1 shows the results derived from the application of the GLCM and Knn feature extraction and analysis routines to the x200 images. The overall rate of classification consistently rises above 90%, and represents a performance level that exceeds the 62% maximum of the original x500 set. Differences in the clarity of the digital image samples at the two distinct levels of magnification has provided fundamental evidence of the damaging effect of the x500 blurred areas on the rate of classification. It has also demonstrated the second order feature's sensitivity to noise and its difficulty in

distinguishing between classes that possess a significant amount of crossover information. Alternatively, the 31% difference in classification performance may stem from the fact that x200 set is formed from fewer samples and divided into fewer groups. Further experiments are performed within this chapter to help reach a final decision.

The results of Table 6.1 verify two important points made in Chapter 5 regarding the K-nearest neighbour classifier.

Table 6.1: The effect of varying K, for the x200 GLCM feature measurements

K	% CLASSIFIED FOR EACH OF THE 3 ANNEALING GROUPS			OVERALL % CLASSIFIED
	5.38% Fe	10.25% Fe	13.96% Fe	
2	100	80	93	91
3	100	80	93	91
5	100	80	93	91
7	100	80	93	91
9	100	80	86	88
11	100	93	86	93
13	100	100	80	93
15	100	100	80	93
20	0	0	73	24
25	0	0	40	13

Firstly, the value K can be regarded as a measure of cluster compactness (density) and an estimation of cross-cluster contamination and boundary shape.

Secondly, the order in which the measurements are analysed by the clustering system is (once again) a significant problem. In this case, the optimum images are examined before the over-alloyed, therefore, the optimum class gains the initial advantage by expanding its borders and capturing any incoming over-alloyed data points as its own, as the value of K increases. Therefore, the initial shape and order of cluster expansion are significant

parameters that affect the final classification result. The high rate of classification for the under-alloyed class is a result of the high level of homogeneity and basic intensity and structural differences from the alternative annealing types, its cluster is well separated.

Tables 6.2 and 6.3 illustrate the differences in performance between the 2-D sparse annular and single turn spiral templates of varying size, when applied to the x200 GLCM filtered and unfiltered images. Using the clear x200 samples the morphological process remained unaffected by the problem of information loss, caused by blurred regions limiting the level of useful data extraction (Chapter 5, section 5.3.1, Figure 5.7a). As a result, the classification performance of the two different types of probe for both the filtered and unfiltered images is of a similar rate. However, the two-dimensional structure (Table 6.2) when applied to the unfiltered images, offered fluctuating rates throughout the threshold range and performed poorly at the higher levels. Conversely, even with the absence of any blurred areas, the filtering operation still provided enough openings on the image surface to enable the probe to extract crystal measurements at all levels of intensity. When analysing the filtered images, every flat sparse annular template generated steady results that ranged from a minimum of 72%, to a maximum of 99%. The spread compares favourably with the extended drift of between 30% and 100% for the equivalent non-filtered samples. This emphasised the point made in the previous chapter that the filtered image reduces the probability of producing a weak classification result due to the setting of the template intensity to a non-optimal height.

Table 6.2: Summary of classification for various sized 2-D sparse annular probes, for the x200 image set

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	99	98
20	95	97
40	99	99
60	98	97
80	95	96
100	94	94
120	92	90
140	82	84
160	86	93
180	53	95
200	30	90

(a) "Small nut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	100	98
20	96	99
40	97	98
60	100	99
80	97	98
100	97	94
120	91	94
140	91	86
160	82	82
180	82	84
200	43	90

(b) "Donut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	88	97
20	98	96
40	88	95
60	92	95
80	90	93
100	88	89
120	84	84
140	85	82
160	75	80
180	70	85
200	37	80

(c) "Huge nut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	50	72
20	98	98
40	93	95
60	75	93
80	83	95
100	90	93
120	86	89
140	92	90
160	85	90
180	80	87
200	73	91

(d) "Giant nut" template

The 3-D single turn spirals were unable to match the 100% classification rate produced by the "donut" template (Table 6.3). Nevertheless, it consistently produced rates of up to and including 99% for both filtered and unfiltered x200 images. Without the blurred areas of the x500 samples, the extra level of analysis introduced by varying the height of each element within the spiral structure meant that it was better able to extract the 3-D profile of the valley to peak regions, with or without the filtering routine.

Table 6.3: Summary of classification for various sized 3-D single turn spiral probes, for the x200 image set

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	98	98
25-100	97	97
35-110	95	95
45-120	95	96
55-130	95	95
65-140	92	93
75-150	91	92
85-160	94	91
95-170	86	92
105-180	87	88
115-190	85	87
125-200	85	93

(a) "Small nut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	99	99
25-100	99	99
35-110	99	99
45-120	97	97
55-130	97	97
65-140	96	95
75-150	95	92
85-160	94	93
95-170	91	90
105-180	90	86
115-190	91	89
125-200	88	84

(b) "Donut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	90	97
25-100	93	96
35-110	95	95
45-120	91	95
55-130	91	92
65-140	92	90
75-150	88	90
85-160	88	91
95-170	90	90
105-180	87	87
115-190	86	85
125-200	88	83

(c) "Huge nut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	84	93
25-100	80	95
35-110	85	97
45-120	85	95
55-130	84	95
65-140	85	92
75-150	83	94
85-160	90	93
95-170	89	90
105-180	86	91
115-190	85	90
125-200	89	88

(d) "Giant nut" spiral template

The GLCM probe intensity optimisation procedure required a drastic cut in the number of locus-element (l-e) measurement decision switches (key-ranges) before the rate of classification approached the levels of the original 2-D and 3-D templates (Table 6.4). Tables 6.4a and 6.4b represent the effect of filtering the GLCM to reduce the selection of l-e switches. Comparing the performance against the unfiltered GLCM equivalent

structures, show an improvement of 28% giving a classification rate of 95% as opposed to 67%.

Table 6.4: Summary of classification for the optimised templates, using the locus-element GLCM filtering options 1 and 2 for the x200 images, data set 5 training images/class

X%	% CLASSIFIED
10	66
20	66
30	66
40	77
50	93
60	94
70	95
80	66
90	68

(a) GLCM filtering option 1

Removing elements with frequency of occurrence $< x\%$ of the maximum occurring pixel-pair combination over the entire GLCM

X%	% CLASSIFIED
10	62
20	65
30	61
40	56
50	65
60	65
70	80
80	85
90	94

(b) GLCM filtering option 2

Removing elements with frequency of occurrence $< x\%$ of the maximum occurring pixel-pair combination on a line-by-line basis

Even though the harsher filtering of the first option performed poorly on the original set of x500 images, for the x200 set, its harshness proved useful at removing enough key-ranges to produce the highest rate when using 5 training images/class. The severity of the filtration method based on removing elements less than the overall maximum value (option 1, Table 6.4a) should only be used where the in-class bond is particularly tight, as with the x200 samples (Appendix 5). Remembering (from Chapter 5, section 5.4.3) that filtering the GLCM prior to template intensity extraction shifts the emphasis to in-class similarities rather than between-class differences. The harsher the filtering, the greater the level of in-class homogeneity expected. Unlike the x500 collection, each x200 class

exhibits distinct, undamaged crystal properties that vary greatly between each annealing type. This allows the harsh filtering of the first option to retain only the important, common features within each class, whilst removing the crossover elements. The less intense line-by-line filtering routine (option 2, Table 6.4b) retained a greater amount of crossover data and expected the incoming test image to exhibit the same properties of a specific class over the entire locus-element GLCM feature matrix. This resulted in each optimised template extracting crossover data from the alternative classes, reducing the overall classification rate. Only when the milder line-by-line filtering routine is adjusted to remove a greater number of elements (option 2, Table 6.4b, when X equals 90%), does the process begin to exploit the in-class similarities, resulting in a noticeable increase in classification performance.

Increasing the number of training images for the filtered GLCMs produced a beneficial response from the second method of matrix filtering (Table 6.5). Where each of the five classes of the x500 samples share similar surface characteristics with texture clarity reduced by the depth of field blurring, each of the three x200 classes demonstrate unique, unmistakable crystal formations that help to visually distinguish one class from another. Therefore, by increasing the number of available x200 training images, the amount of effective measurement key-ranges for each template prototype also increases. Unlike the x500 equivalent, where the larger ranges increased the possibility of cross-class probe measurements (Tables 5.8, 5.11 and 5.12, Chapter 5), the in-class cohesion of the x200 set ensures that there is only a minimal risk of intensity contamination at the higher training numbers. The results for both filtering methods managed to better the maximum rate of 67% for the unfiltered equivalent, with the biggest improvements realised when the maximum filtering percentage value, was selected on a line-by-line basis (option 2,

Table 6.5b). Using the second GLCM filtering option removed less information from each matrix. This is highly desirable in cases where there are a large number of training images, since there is a higher probability of the in-class training GLCMs resembling one another and representing the full range of possibilities within the annealing group. Therefore, the lighter filtering option enabled the selection routine to produce matching key-regions that maximise the available locus and element measurement ranges.

Table 6.5: Varying the number of training images for template optimisation scheme

NUMBER OF TRAINING IMAGES	% CLASSIFIED
1	33
3	75
5	95
10	92
20	91

(a) GLCM filtering option 1

Removing elements with frequency of occurrence $<x\%$ of the maximum occurring pixel-pair combination over the entire GLCM

NUMBER OF TRAINING IMAGES	% CLASSIFIED
1	33
3	66
5	94
10	98
20	95

(b) GLCM filtering option 2

Removing elements with frequency of occurrence $<x\%$ of the maximum occurring pixel-pair combination on a line-by-line basis

At 98% (option 2, Table 6.5b) the 10 images/class result offers a slight improvement over the previous best of 95% (5 images/class) and improves upon the 93% maximum of the GLCM and Knn routines. Despite this increase, the optimisation technique still failed to

match the classification performance of the original 2-D and 3-D structures (100% and 99% respectively).

For the Brodatz subset, the GLCM and Knn approach at best classified only 50 out of the available 162 images, giving a peak classification rate of 31%. The poor performance stems from a number of possibilities.

Firstly, the use of non-optimal values for the GLCM parameters, distance d and angle θ, generate a feature matrix that introduces the problem of information loss [88, 149]. The whole purpose of performing feature analysis rather than image analysis is to quicken the process and improve the rate of classification by ignoring redundant data. However, if the data extracted does not explicitly describe the important structural formations on the image surface then the feature has failed to perform the required task. Nevertheless, experiments in varying the angle and distance parameters on the Brodatz images failed to show the same variations in the rate of classification as the investigation into mortar densities in [118], with results only showing an improvement of 2%, giving 33% overall. This is partly due to the choice of classifier and the GLCM descriptors analysed, in addition to the high number of classes and texel variations within the Brodatz subset, compared to the relatively trivial number brick textures analysed in [118].

For the Brodatz image subset, the second possibility for the poor GLCM and Knn classification result, is the likelihood that too many texture measurements were used to define each image. With ten dimensions extracted from each GLCM, there is an increased chance of the measurements offering conflicting results. However, removing certain dimensions to optimise the feature set would only offer a limited improvement in

separation, since no one descriptor was enough to completely classify any single texture type (due to the amount of class overlap).

The final possibility is that the Knn classifier required more training images to increase the density of the prototype clusters and offer a more comprehensive and detailed representation of each class of texture. Extracting yet more 128^2 pixel regions from the same Brodatz subset would not fully test the feature extraction and classification system. The limited resolution would only allow for the extraction of overlapping segments. A difference of one or two pixels between one sub-image and the next can not hope to match the utility of a completely new sample. For example, the image in Figure 6.1a shows a collection of bean objects. By extracting a number of overlapping sub-images at different co-ordinates within the same image, various viewpoints of the same unit pattern is presented. Whilst this increases the number of classification candidates, the amount of material that crosses-over restricts the usefulness of each sample as it fails to represent the possible range of bean variations within the class. Furthermore, by ignoring how a bean interacts with the surrounding environment and its neighbouring bean objects, the classification results would give a false impression on how the technique would perform within a live setting. Without separate sub-images to fully test the capabilities of the feature extraction and classification system, no overlapping image segments were extracted to increase the Knn cluster densities and to examine the above theory.

Tables 6.6 and 6.7 represent the performance differences between the 2-D and 3-D templates when applied to the standard non-filtered and GLCM filtered Brodatz samples. The R-C operation when using the sparse annular "Donut" template, managed to classify

at best over 122 out of the original 162 Brodatz samples, to produce the highest classification rate of 76%.

Table 6.6: Summary of classification for various sized 2-D sparse annular probes, for the Brodatz image subset

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	70	67
20	72	68
40	64	73
60	66	65
80	58	64
100	37	46
120	20	40
140	9	29
160	3	27
180	2	16
200	0	2

(a) "Small nut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	61	61
20	75	72
40	62	76
60	67	76
80	61	67
100	54	61
120	40	45
140	29	44
160	14	24
180	3	14
200	0	2

(b) "Donut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	29	29
20	54	52
40	53	60
60	56	61
80	54	65
100	51	62
120	49	42
140	38	35
160	24	30
180	9	16
200	4	8

(c) "Huge nut" template

TEMPLATE INTENSITY	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
0	15	23
20	57	53
40	60	63
60	55	64
80	58	59
100	63	58
120	51	47
140	45	46
160	38	30
180	27	30
200	17	19

(d) "Giant nut" template

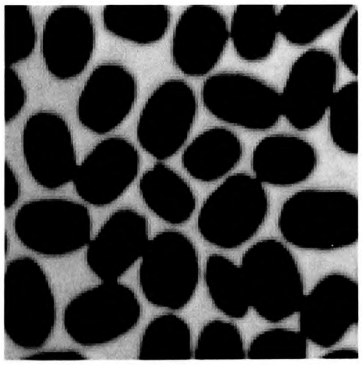
Confirmation of the ability of the GLCM filtering procedure to improve template access for all forms of natural texture is presented by the fact that the highest result was produced by the filtered images (Table 6.6b). As with the x200 image set, the differences between the results of the filtered and unfiltered versions remained similar throughout the

template intensity range. Without any trace of the blurred areas that affected x500 samples, the filtering routine only offered an improvement of 1% over the non-filtered peak rate of 75%. However, where the unfiltered results fluctuate and rapidly fall, the pre-processed equivalent steadily drops right to the template intensity limit of 200.

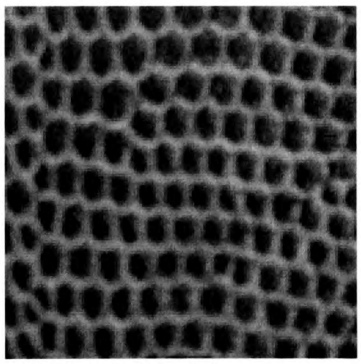
With a wider variety of definite texel patterns and object boundary shapes, the size of the probe becomes a more significant factor when analysing the Brodatz collection. The sparse annular structures perform at their best when they are forced to examine local texture variations rather than the wider unrelated pixels. The larger sized "huge nut" and "giant nut" probes were unable to analyse and contain the important unit patterns and performed the poorest when the smaller templates were reaching their peak result. Nevertheless, the careful choice of the template threshold level still remains the overall key factor in trying to classify the greatest number of images. The optimum intensity range for the filtered images was between 40 and 80 depending on the size of the applied structure. For the non-filtered samples, the highest classification results were generated at the lower threshold levels of around 20 for the smaller templates, and exactly 60 and 100 for both the "huge nut" and "giant nut" respectively. For the "small nut" and "donut" probes, the small increase in threshold intensity from the unfiltered 20 to the filtered 40, coupled with only a slight improvement in the rate of classification, indicate that the GLCM filtering routine is dropping the level of only a limited amount of the higher frequently occurring pixels. The extraction of a histogram defining the number of unique grey-tones that form each examined Brodatz sub-image, indicated that around 55 unique intensities form each sub-sample. Consequently, the pixel intensity level within the 25^2 regions of filtering analysis is dominated by a select number of individuals that tend to lie around the same value (providing the window contains primarily elements of the main

unit pattern). This restricts the effect of the filtering process and preserves the main texel boundaries. The natural formations of the Brodatz sub-images already contain a significant proportion of low-intensity pixels, which allow the analysis structure to better examine high contrast areas. Consequently, the filtering process only offers a slight improvement in template access for these particular images (Figure 6.2). For the remaining samples however, by removing the more frequently occurring regional elements whilst retaining basic boundary shape, and adjusting the height of the probe to compensate (in this case by 20), the template is able to measure the image to a higher standard and offer an overall improved rate of classification (Figure 6.3).

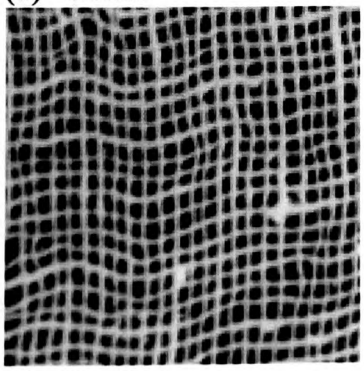
Figure 6.2: Selection of images from the Brodatz set that contain areas of high contrast (and therefore do not benefit greatly from the GLCM filtering process)



(a) Beans

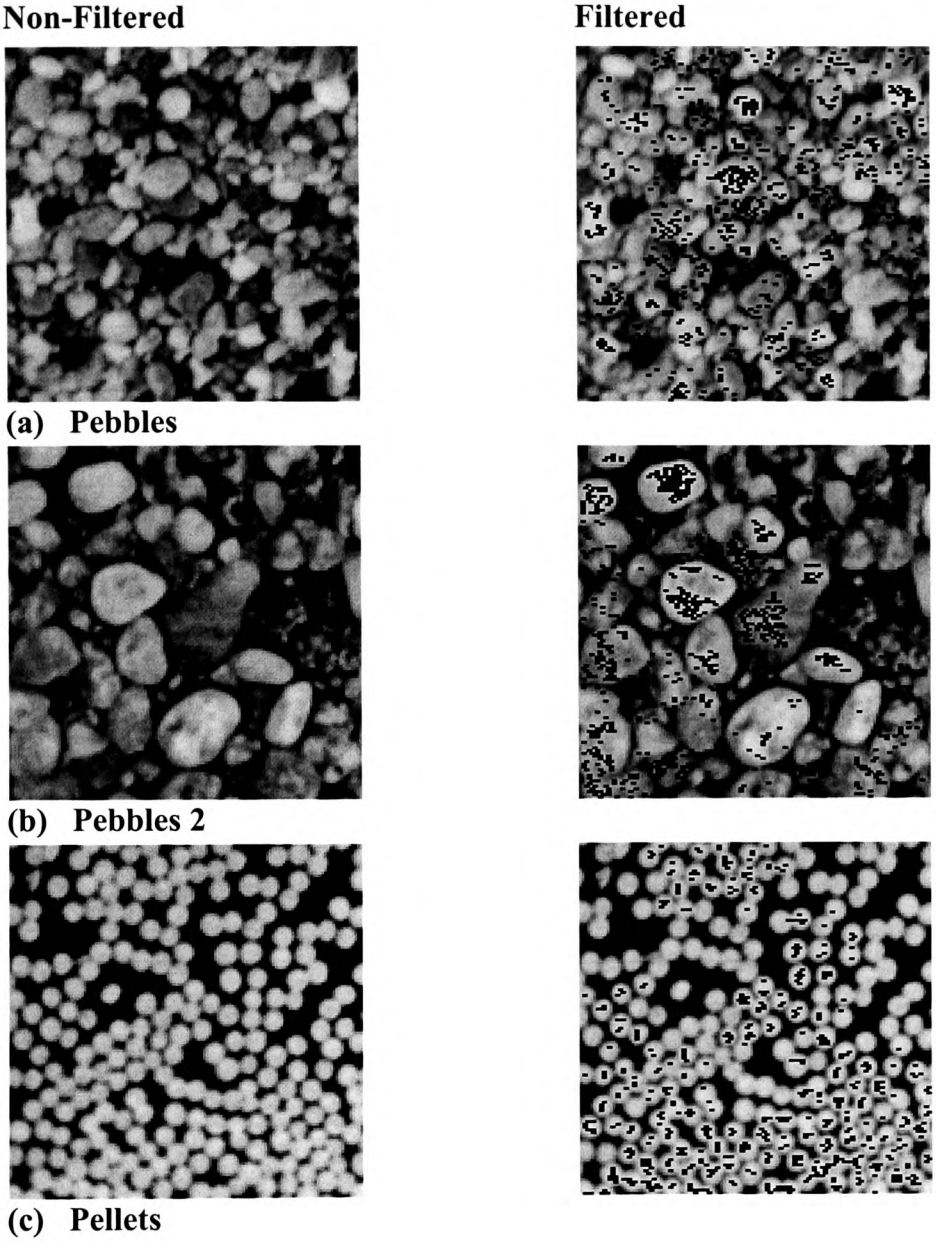


(b) Reptile skin



(c) Burlap

Figure 6.3: Selection of images from the Brodatz set that benefit from GLCM filtering



The single turn spiral results of Table 6.7 demonstrate that the 3-D structure is unable to match the performance of its 2-D counterpart. The "small nut" single turn template produced the highest overall spiral classification rate of 72%.

Table 6.7: Summary of classification for various sized 3-D single turn spiral probes, for the Brodatz image subset

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	62	72
25-100	60	72
35-110	53	68
45-120	48	62
55-130	40	50
65-140	26	47
75-150	22	44
85-160	20	37
95-170	20	33
105-180	16	34
115-190	13	24
125-200	6	27

(a) "Small nut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	70	69
25-100	59	69
35-110	59	70
45-120	59	70
55-130	53	67
65-140	51	63
75-150	49	58
85-160	49	58
95-170	45	56
105-180	36	50
115-190	30	41
125-200	17	33

(b) "Donut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	56	57
25-100	52	59
35-110	56	56
45-120	58	61
55-130	53	60
65-140	50	62
75-150	49	59
85-160	50	55
95-170	46	56
105-180	39	48
115-190	45	37
125-200	33	35

(c) "Huge nut" spiral template

TEMPLATE INTENSITY RANGE	% CLASS, NON-FILTERED IMAGES	% CLASS, FILTERED IMAGES
15-90	50	56
25-100	45	54
35-110	40	58
45-120	45	57
55-130	51	56
65-140	43	58
75-150	48	59
85-160	56	51
95-170	48	52
105-180	50	43
115-190	51	40
125-200	45	34

(d) "Giant nut" spiral template

For certain images, the full containment of the analysis probe is highly desirable and is one area that the 2-D sparse annular structure has the advantage. Using the beans image as an example (Figure 6.2a), without the increasing height of the spiral design the flat template has a greater possibility of being fully contained at the point where it approaches the high intensity pixels, from either within the darker bean objects or from the GLCM filtered regions. The ability to fully contain the geometric shape at certain points within

the image, generates a unique histogram feature that captures any unusual texel properties that may not exist in the alternative classes. The spiral probe on the other hand, tends to either partially contain or completely ignore certain pixel formations because the grey-scale range is so narrow. By limiting the level of containment, the histogram fails to isolate the important classifiable data. This leads to a feature that is not entirely accurate and one that yields a similar set of texture measurements for each image class. The single turn spiral structure is unable to represent images that contain only a low number of grey-tone variations to the same precise level as the sparse annular templates. Moreover, without the presence of any blurred regions, the risk of fully containing any unwanted crossover information by the 2-D structure is no longer a serious issue. Consequently, the three-dimensional model is placed at a secondary position when dealing with alternative forms of natural texture, where the 3-D profile is not as important as the two-dimensional shape of any of the significant texel primitives.

Attempting to generate a series of optimised templates for the Brodatz image subset, revealed a further limitation of the GLCM probe threshold optimisation scheme. With only around 55 unique grey-scale intensities for each individual sub-image group and 18 different types of texture, the number of potential locus-element measurement decision switches is severely limited to at best 2-3 per class. This also depends on the strict selection criteria defined in the previous chapter, i.e. a measurement switch is only assigned to a class optimised structure, if it either exists uniquely within the training group, or exists at a higher frequency of occurrence than the alternatives. The problem now becomes more significant since the analysis template for certain classes may never extract a single measurement from its own group while other strong individuals may dominate the classification arena. Additionally, with so few morphological measurements

taken per image, the effect of crossover intensities within the test data become more significant, leading to a probable increase in the number of misclassified images.

The following experiment attempts to complement the previous set of results and lead to the conclusion that the use of blurred images for training and testing is the main reason for the underwhelming x500 classification rates, rather than the operating limitations of the procedures investigated. It also offers an indication of whether the high x200 classification results for all of the applied techniques, is a consequence of only examining a limited number of images and annealing types.

Three individual groups of 120 images were drawn from the original x500 galvanneal samples and separated into their respective annealing classes. This provides an opportunity to assess the utility of the original sample set when arranged in a similar configuration to the x200 collection, i.e. 3 phases of annealing, each with 40 images. Analysing the new arrangement presents a clear estimate on the capacity of the blurred regions to affect the accuracy of the morphological extraction and classification approaches, by comparing the performance with those of the original x500 and x200 sets.

Table 6.8: Summary of classification for 3 subsets of 120 unfiltered x500 images, using only three classes of annealing, to analyse the effect of class and image count on the rate of classification

APPLIED TEMPLATE	% CLASSIFIED SUBSET 1	% CLASSIFIED SUBSET 2	% CLASSIFIED SUBSET 3	% CLASSIFIED FULL SET
"Giant nut", intensity 40	83	82	85	67
"Huge nut", intensity 20	68	72	75	66
"Donut", intensity 20	70	80	78	71
"Small nut", intensity 20	71	75	77	71

Table 6.8 depicts the classification rates of the 3 subsets of 120 x500 galvanneal samples and compares each percentage figure against the complete set of 395. Four templates of varying size were applied to each individual subset, with their threshold level determined by selecting the value that produced the highest results on the complete collection of x500 unfiltered samples.

For the "huge nut", "donut" and "small nut" probes, the classification rates differ to that of the "full set" percentage by only a small degree. With a maximum difference of only 9% produced by the larger "huge nut" and "donut" structures, the initial belief is that the depth of field blur is still at fault, and the reduction in the amount of information analysed is of limited benefit. However, the "giant nut" probe offers a perspective that contradicts the view presented by the alternative templates. For every single subset of 120 random x500 images, the largest applied structure offers an increase of at least 15%. Rather than introducing an element of doubt, the results serve only to verify a number of fundamental points.

Due to the magnitude of its construction, the "giant nut" probe is unable to examine any local changes in crystal shape and size. The pixels it encounters represent the grey-tone values that are not strongly bounded by the local intensity variations of the crystal formations. Without the ability to contain any of the important regional structures, the template is at a disadvantage when attempting to classify any form of natural texture, where the detection of any fine detail is paramount. Hence, the poor classification results for every type of image analysed by the "giant nut". Even though the obscure areas within the full set of x500 galvanneal samples contain classifiable data, if a template is not containing any of the prime crystal detail, then it is only analysing and extracting

damaged, crossover information. By removing a number of blurred images and annealing phases, the probability of the "giant nut" template containing class-common, substandard data declines greatly. The risk is further reduced, by restricting the test data to only images from the optimum-alloyed, extreme over and extreme under-alloyed collections, where the chance of similar structures and intensity levels occurring is not as great as in the neighbouring groups. The result is an improvement in the rate of classification for the "giant nut" template, but only through luck rather than judgement.

6.2.3 Conclusion and Discussion

For the applied techniques, the rate of classification dramatically increased when the x200 galvanneal steel images were being analysed. Both the morphological and second order statistical approaches demonstrated an increase in performance over the x500 samples by around 20% and 30% respectively.

Although Table 6.8 places some added confidence in the previous set of results and substantiates the statement that, the correct choice has been made on the selection of a feature extraction and analysis routine. It shifts the problem to an alternative decision between the weak x500 classification results due to the depth of field blur, or performance enhancements purely because of a decrease in image and class-count. To reach a final conclusion, a few more points need mentioning to supplement those made in the latter part of the previous section. Referring back to the experiments involving the x200 samples and the reduced x500 data set of Table 6.8, a number of items stand out that places the blame squarely on the blurred data.

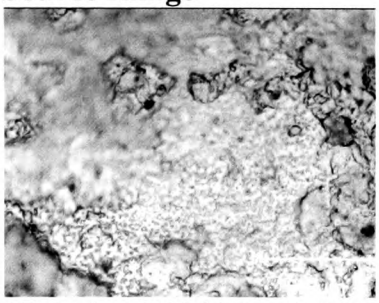
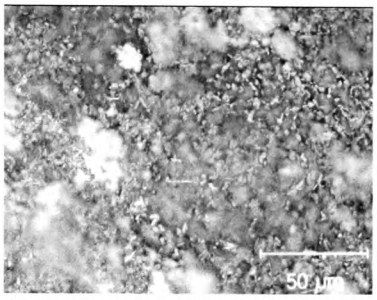
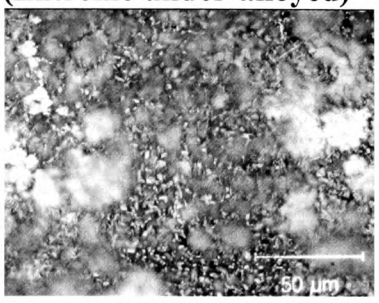
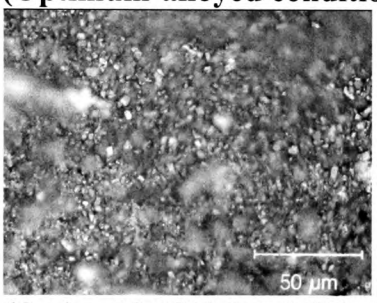
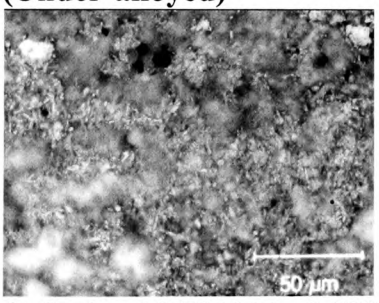
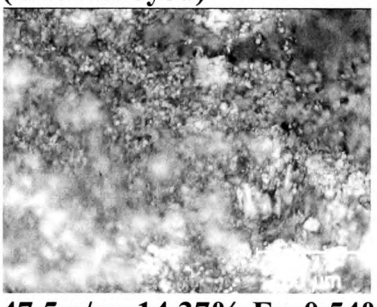
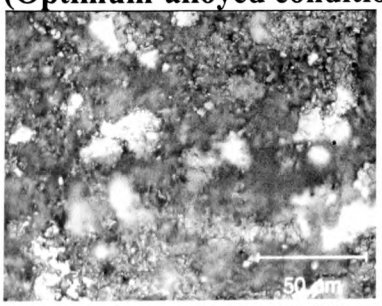
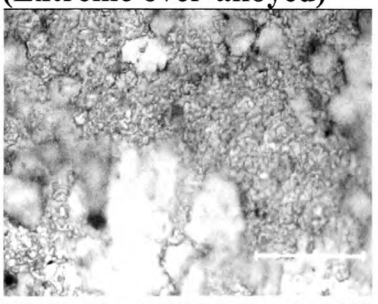
Firstly, if the reduced number of classes had helped to produce the 100% x200 classification result, then the increased number of texture types within the Brodatz collection should have in theory, produced an abysmal set of percentages for the standard morphological approaches instead of the respectable maximum value of 76%. Whilst for the GLCM and Knn procedures, the Brodatz result of 31% does represent an abysmal mark and a dramatic drop in performance from the x500 and x200 images, the reasons for the drop has already been discussed and ultimately has no bearing on the final outcome.

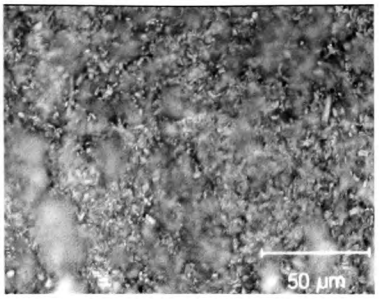
Secondly, if the blurred regions of the complete x500 image set were not making such a noticeable impact on the standard of analysis and classification, then the results of the spiral probe should vary between the filtered and unfiltered x200 images (Table 6.3). The results indicate that the blurred areas prevalent at the highest magnification were affecting the level of useful data extraction across the entire image, and were interfering with the standard of containment for the clear crystal formations.

As a final condemnation of the low quality x500 galvanneal steel samples consider the images of Figure 6.4.

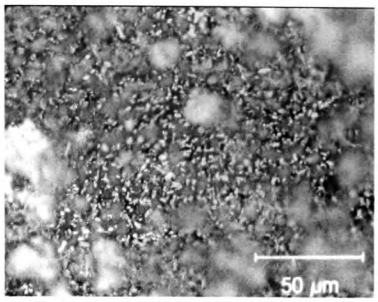
Based on the results generated by the "huge nut" template (all structure sizes reached the same conclusion), the following figures represent a small selection of the images wrongly assigned to the specified class. The structure's intensity was fixed at the zero level in order to focus on crystal formation size and to analyse plateau regions, which include the cloudy texel arrangements. In every case, the assigned image/class is a result of the extraction, analysis and containment of the blurred crossover information rather than any perceived similarities in crystal size between the groups.

Figure 6.4: **Samples from various annealing classes that are considered to possess the same structural qualities due to extensive image blurring**

Source Image	Classified as
	
37.6 g/m, 5.38% Fe, 0.66% Al (Extreme under-alloyed)	**41.9 g/m, 10.79% Fe, 0.59% Al (Optimum-alloyed condition)**
	
41 g/m, 8.42% Fe, 0.61% Al (Under-alloyed)	**43 g/m, 12.08% Fe, 0.6% Al (Over-alloyed)**
	
41.9 g/m, 10.79% Fe, 0.59% Al (Optimum-alloyed condition)	**47.5 g/m, 14.37% Fe, 0.54% Al (Extreme over-alloyed)**
	
43 g/m, 12.08% Fe, 0.6% Al (Over-alloyed)	**37.6 g/m, 5.38% Fe, 0.66% Al (Extreme under-alloyed)**



47.5 g/m, 14.37% Fe, 0.54% Al (Extreme over-alloyed)



41 g/m, 8.42% Fe, 0.61% Al (Under-alloyed)

This indicates that the blurred x500 samples are to blame for the poor classification results, rather than the number of images/texture types, or the ability of the applied techniques.

6.3 Interim Conclusion

Applying the morphological and GLCM approaches to both the x200 galvanneal images and the alternative textures of the Brodatz collection, has produced a set of results that support the conclusions drawn from the original x500 work, and has assisted in establishing the final set of conclusions:

(i) The quality of the source data is paramount. Poor texture resolution aggravates the conversion from digital image to feature representative. Thus, reducing the feature's ability to accurately define the original sample, irrespective of the applied texture extraction and analysis technique.

(ii) For the galvanneal data set examined, image clarity is more important than fine crystal detail providing the fundamental class differences are prevalent at the chosen scale. It can be stated with some confidence that the x200 galvanneal steel

samples produced the higher classification results due to the absence of any obscure texture regions, rather than the reduced image and class count.

(iii) The suitability of a particular morphological template for feature extraction and analysis is based upon the type of texture analysed. Both 3-D spiral and 2-D sparse annular probes have their benefits. The spiral for analysing 3-D texture profile and the 2-D sparse annular for extracting a greater amount of information from images with a narrow grey-scale range.

(iv) Resetting the pixel intensity to zero at strategic points within an image using the grey level co-occurrence filtration routine, increases the level of accuracy for the applied morphological template whilst retaining basic unit pattern boundary shape.

(v) For images formed from complex texture samples, a system that evaluates both the fundamental elements and the finer structural details performs to a higher standard, to one that only concentrates on a single area. I.e. probability density function and grey level co-occurrence matrix versus the rank conditioned-morphological transform.

Chapter 7

Summary and Suggestions for Future Work

7.1 Introduction

The investigated image processing techniques have proved to be capable performers for the analysis and classification of galvanneal steel and other forms of natural texture. Nevertheless, they each possess a number of weaknesses that have become apparent as they have been applied to the steel strip surface. These limitations have provided a clear direction for future study, involving the analysis and classification approaches that have repeatedly classified the annealing phase of the galvanneal samples to a high standard.

Chapter outline:

(i) Summary of Limitations for the Techniques Investigated, as Applied to the Galvanneal Steel Image Surface - Based on the results produced by the x500 and x200 galvanneal image sets, the summary highlights any weak points within the applied techniques as a prelude to the following section.

(ii) Further Work and Suggested Improvements - Focussing on the procedures that offer the greatest scope for improvement, a number of ideas are presented that may be implemented over both the short and long term periods.

7.2 Summary of Limitations for the Techniques Investigated, as Applied to the Galvanneal Steel Image Surface

7.2.1 Wavelet Transform

A compression routine based on the wavelet transform was applied to the galvanneal steel samples in an attempt to isolate the crystal boundary shapes from the surrounding blurred texel regions (Chapter 4, section 4.2). These crystal boundaries could then be extracted and measured to help form the bases of a classification feature set.

The routine failed to segregate any significant structures and only managed to increase the area of blurred information. By removing the lower frequency components within the wavelet domain, the compression scheme was actively targeting and destroying the very features it was trying to enhance. This resulted in a lower quality image, stripped of its class-defining crystal detail.

7.2.2 First Order Statistical Pattern Recognition

The first order probability density function (PDF) was introduced as a feature extraction technique that could exploit the fundamental pixel intensity differences between each annealing class (Chapter 4, section 4.3). By extracting the four moments defined by

Levine [113], together with an additional entropy measurement, it was expected that this simplistic method of image analysis would offer a quick and efficient solution.

Analysing images of the same galvanneal type, but captured using alternative light sources, pinpointed a crucial failing of the one-dimensional histogram. The first order measurement defines the probability of a grey-scale pixel existing within the sample image. Not only does this generate a feature that ignores important crystal boundary shape, it also exposes the PDF as being overly sensitive to the value of an individual pixel. These are undesirable attributes that fail to exhibit the robust and reliable properties required for the approach to be used in the harsh, dusty environment of a steel strip galvannealing line.

7.2.3 Second Order Statistical Pattern Recognition

The utilisation of the grey level co-occurrence matrix (GLCM) as a feature for steel texture extraction and analysis is based on its ability to analyse the crystal structure formations rather than basic grey-tone value (Chapter 4, section 4.4).

Besides the obvious increase in computational time in direct relation to the number of grey-scale values present in the source material, the GLCM as with the first order analysis technique, is noticeably affected by the value of a single pixel. Whilst not to the same extent as the first order histogram, it still exhibits differences that have a noticeable effect on the rate of classification. If the variations in pixel intensity is non-localised and spread throughout the entire image, then irrespective of the fact that the feature relies upon the interactive nature of pixel-pair values rather than an individual point, the spread of frequencies will still drift from the expected bias locations.

Additionally, if there is not enough grey-tone variation within the source data the measurements extracted do not offer enough scope for clear separation, especially where the class-count is high e.g. the examined Brodatz subset. This places further pressure on the classifier to isolate the matching class from the merging cluster boundaries, an operation which more often than not relies upon the density of a cluster rather than the level of group separation. This mode of operation tends to produce unreliable results (Chapter 4, section 4.5.5) as the in-class clusters approach their density limit.

7.2.4 Genetic Algorithms

The genetic algorithm (GA) was introduced in an attempt to examine the performance of the grey level co-occurrence matrix without the benefit of a priori knowledge (Chapter 4, section 4.5.1). A GA search procedure was used to generate a population of weighting values, that could be used to draw in each of the ten descriptors based on the GLCM to a measurement and class specific centroid.

Applying the search routine to the co-occurrence feature measurements demonstrated a fundamental problem with using the genetic algorithm as a method to reduce cluster size. If the incoming test data generates clusters that are widely spread and illustrate a high level of between class crossover, then the GA is unable to produce a set of weighting values that are capable of reducing the in-class variance whilst correcting any cross-cluster contamination.

7.2.5 K-Nearest Neighbour

The K-nearest neighbour classification (Knn) scheme was employed as an alternative classifier to the genetic algorithm search procedure (Chapter 4, section 4.5.4).

This alternative approach to galvanneal classification failed to offer satisfactory results for two very important reasons.

Firstly, the classifier displayed sensitivity towards the order in which it encounters the five annealing types, favouring the earliest cluster over the others.

Secondly, the size and shape of the original training clusters affects the overall assignment of the incoming test data. These factors have a profound effect on the Knn's performance and place a strict limit on its ability to achieve respectable classification results on a live galvannealing line.

7.2.6 The Rank-Conditioned Morphological Transform

The rank-conditioned (R-C) morphological transform was applied to the steel data in an attempt to extract a more accurate feature measurement that could better define the five annealing states than the alternative techniques (Chapter 5, section 5.2).

Whilst the R-C operation produced some of the highest rates of classification irrespective of the texture contained within the source material, the method still has a number of issues that require attention.

Firstly, the geometric template used for analysis remains a rigid structure that is unable to accurately represent the variations in the surface texture topology.

Secondly, the accuracy of the morphological measurement lowers as the clarity of the steel crystal structures is reduced through the problem of depth of field blur.

Thirdly, a pixel point is either contained or missed by the analysis probe, there is no indication of the level of containment or the distance between the structure and the target grey-tone, giving a less accurate feature representation of the original sample.

Finally, the threshold level and to a lesser extent, the size of the R-C template is a major point of concern. By incorrectly setting the two parameters, the probe generated is sub-optimal and out of phase with the surface structures, i.e. there exists an optimum size, shape and height combination for all types of texture, using a poor combination serves only to generate classification rates that do not fulfil the technique's potential.

7.2.7 Image Filtering using the Grey level Co-occurrence Matrix and the Rank-Conditioned Morphological Transform

The GLCM image filtering routine was developed as a technique to address the problem of template "skimming" caused by the blurred regions within the x500 samples (Chapter 5, section 5.3). It also increases the level of template access throughout the texture matrix for other types of image whilst preserving basic texel shape, and thus enables a more accurate feature representation of the original data.

The utility of the image filtering procedure depends entirely on the type of texture formations present within the test sample. If the sample contains a limited number of domineering grey-scale pixels then a greater amount of data is removed. Conversely, an image with grey-tones evenly spread produces a fine level of localised filtering. This mode of operation gives the routine an unpredictable nature that restricts its ability to operate successfully without adjustment on images where the between class structural formations vary greatly.

7.2.8 Probe Parameter Optimisation Scheme

The probe parameter optimisation scheme was introduced as a technique that could remove the guesswork from template intensity selection and to always ensure that only class optimal information is analysed (Chapter 5, section 5.4). The technique would address some of the problems associated with the original R-C procedure (section 7.2.6), and in theory limit the number of template re-applications required to obtain the optimum result.

The creation of a process that is able to generate the most efficient set of threshold values for any shape or size of template and for any classification task, is highly desirable. However, the process in its present form is unable to compete with the original morphological method for a number of reasons.

Firstly, the scheme was designed to restrict the analysis window and focus on important crystal formations within each group, rather than recreate the undesirable blanket examination offered by the standard morphological technique. Whilst this reduces computational overheads, it produces a system that is significantly affected by noise and

cross-class grey-scale intensities. In high class count situations or where the grey scale range is restricted, only a limited amount of locus-element (l-e) keys per class are produced. Any amount of intensity crossover therefore significantly increases the probability of an incorrect class assignment. This probability is further increased as the trivial amount of "good" measurements extracted fail to override the decision in favour of the correct group.

Secondly, as the class-count increases so the number of unique or class-defining pixel intensity variations decreases (within a finite grey-tone palette). This has a profound effect on the number of available locus-element keys, with certain classes unable to extract a single measurement while others dominate.

Similarly, utilising training sets that use a severely restricted grey-tone palette limits the amount of l-e keys available and their effectiveness. Even though the keys extracted from the training images may be the ideal collection for the basic subset, their application to the entire group would increase the likelihood of extracting a crossover measurement due to the limited number of unique grey-tones within each class.

Finally, the probe parameter optimisation scheme is unable to compete with an expertly designed template structure. Further work is required on the key-range selection routine, the GLCM filtration scheme and the updated classifier before the system can fulfil its potential.

7.3 Further Work and Suggested Improvements

Focussing on the area of galvanneal steel classification, it is important to take stock of the key issues and classification achievements from the previous chapters, before suggesting how any possible improvements can be examined and implemented.

Up to this point, the rank-conditioned morphological technique has established itself as a credible performer. On the two sets of galvanneal samples and the alternative natural texture sets of the Brodatz collection, it generated respectable classification rates that far exceed the performance of the alternative procedures. Therefore, the "short-term" improvements of 7.3.1 focus on the morphological technique, the various geometric templates it uses and the quality of the galvanneal images it analyses.

Even though the probe parameter optimisation scheme did not perform as expected, it is the one technique that possibly offers the greatest scope for improvement. The "long-term" developments of section 7.3.2 verify the need for improved selection and classification functions, and propose an alternative method of assessing template coverage.

7.3.1 Short-Term

The short-term suggestions represent ideas that could be implemented immediately and instantly provide a possible improvement in classification over the previous best of 80%. The three suggestions focus on two main areas, the source data and spiral adjustment for the standard rank-conditioned morphological transform:

The original x500 collection offered five phases of annealing and approximately 400 individual images. The images provided a detailed view of the galvanneal crystal formations but at a reduced level of image clarity and hence utility. Conversely, the x200 images only offered a limited group of 120 samples and 3 annealing stages but avoided the problem of depth of field blur. Therefore, two additional sets of x200 and x500 samples are required, each divided into 5 complete sets to enable direct comparisons of their classification performance when analysed using the morphological transform. It is also desirable that the x500 images be captured using up-to-date optical equipment to minimise the risk of fine detail blur.

The analysis structures selected for the five spiral voting system managed to produce an 80% classification rate for the x500 images. Further improvements should be gained by fine tuning the size and intensity level of each applied template, rather than combining consecutive intensity sequences of the same sized structure.

The spiral templates introduced in the previous chapter consisted of linearly increasing element threshold values with a fixed stepping value. Whilst it offered an effective additional contrast measurement to better define the 3-D texel profile than the 2-D sparse annular equivalent, its uniform build prevents the reliable extraction of significant class-defining structures. By varying the stepping value at particular points within the probe and increasing the number of turns, it is able to evaluate finer intensity changes within the image and determine minute variances within the crystal profile at the predefined template segments. Further adjustment of the rate of spiral compression at various positions could enable each segment to be tailored to suit each individual class in an attempt to increase the sensitivity of the probe to any changes in annealing characteristics.

In a basic experiment, the technique has already shown an improvement in the rate of classification for an individual template. Using only a single turn "huge nut" spiral of varying intensity and fluctuating compression levels, the result matched the 80% of the five spiral voting system. A similar result was produced in a trivial experiment involving the rotation of the leading edge of a "donut" spiral probe through the four cardinal points of a compass. The results indicated that whilst direction did not have any noticeable effect on the level of performance, it still managed to classify six more images when the largest threshold values were at the southern most point, producing a new highest classification rate of 81%.

The accuracy of the standard R-C procedure could be considerably enhanced by adding a further dimension to the morphological feature histogram that incorporates a measurement that defines pixel containment in terms of rotational position within the applied template. This addition would help to pinpoint areas that mainly consist of important crystal arrangements, and highlight any significant containment points as the template traverses the image. By combining the idea with a filled template structure, the histogram is able to provide a detailed object map of the sample image.

7.3.2 Long-Term

The long-term plans can be considered as loose concepts based on variations of the probe parameter optimisation scheme. These suggestions require further consideration prior to their development and application.

With the exception of the x200 image set, the template threshold selection procedure in its present form fails to perform to the same standard as the original morphological

technique. The prototype locus-element key-range extraction routine, expects a level of class uniqueness that does not exist within certain classification problems. Therefore, as an alternative to changing the operation of the existing selection function, a basic hybrid texture analysis approach that combines the original and novel morphological techniques is considered.

Using a standard analysis template of a predetermined, non-optimised height, the complete structure (not reduced to l-e pairs) is restricted to only analyse areas defined by a more relaxed set of intensity range values, i.e. a weakly defined set of locus ranges instead of the precise locus-element measurement decision switches. Classification is based on the nearest-neighbour method rather than the updated version outlined in Chapter 5, section 5.4.3.

By changing the process of one optimised probe per class and returning back to the rigid intensity design with the introduction of a locus-range, the analysis window is expanded under a controlled environment. This weakens the effect of any crossover extractions and allows the structure to focus on the essential crystal formations.

An alternative approach and one that maintains the same fundamental ideas as the existing probe parameter optimisation scheme in Chapter 5, is to remove the weak points in the original method and build upon the remaining framework.

In its general form, the concept removes the current system's reliance on individual grey-scale values and focuses on grey-tone intensity difference. Calculating a histogram based on the difference between pixel-pair intensities (grey-level difference matrix [110]) using

the locus-elements pairs of the morphological template as distance and angle parameters generates an outline of key-ranges based on grey-tone variance. The key-range values now determine whether or not a measurement should be taken by analysing the distance between the pixel-pair intensities. If the distance is deemed as containing important class-defining information, then a measurement is extracted; otherwise it is ignored.

The opportunity is available to either use a single fixed template for the entire set of images, in which case each class should maintain a separate library of pixel-pair difference keys, or a collection of probes individually tailored to specific groups. The latter option requires a detailed plan on the method of probe intensity optimisation. One possibility is to use the difference calculated as the template threshold, but more consideration is required.

Using pixel-pair intensity distances as a basis for measurement decisions partially removes the constraint imposed by directly using the grey-scale values. Additionally, the method would be capable of examining image sets with a high class-count and/or a restricted palette unlike the existing optimisation routine.

7.4 Interim Conclusion

The limitations of the applied techniques when assessing the annealing phase of the steel substrate (outlined in section 7.2) offer a natural progress path for future improvements in galvanneal assessment and analysis (section 7.3). The ideas presented also offer a potential route for improvements in classification performance for all types of natural texture.

Chapter 8

Conclusion

8.1 Achieving the Primary and Secondary Goals

The primary objective of the project was to determine the link between the annealing phase of the steel strip surface of actual galvanneal steel coils and a captured digital image equivalent. Early experiments involving wavelet compression and the first order statistical pattern recognition technique, failed to produce satisfactory results. Similarly, the grey level co-occurrence approach whilst producing an a priori classification rate of 100%, was unable to perform to a similar level when combined with the non-a priori genetic algorithm or K-nearest neighbour classifiers, achieving a less than adequate 62%. However, applying the rank-conditioned mathematical morphological procedures of Chapter 5, in conjunction with the image filtering process using the grey level co-occurrence feature to both sets of galvanneal images, offers strong evidence that the annealing state of the steel coil can be established by analysis of the captured samples. Furthermore, the experiments of Chapter 6 confirm that the perceived relationship is chiefly determined by the quality of the conversion from the material steel sample to its digital representation. Basic image quality must be of a sufficient level as to permit feature extraction and hence classification. As a case in point, the rate of classification for

the x200 image set, which remained unaffected by any blurred texture regions, returned a perfect rate of 100%. Conversely, the lower quality x500 collection could only manage at best, 80% (ignoring the 81% derived from the experimental work of Chapter 7). Nevertheless, this still represents a respectable percentage considering that the majority of images had lost between 70% and 97% (Chapter 3, Figure 3.2) of the available fine crystal detail that was present in the original material specimen.

Based on the previous set of results, the morphological techniques have allowed analysis of the steel substrate from the captured version. For the image data under test, the primary goal has been achieved. Further samples are required before complete confidence can be placed in the methods developed before application in the live environment can be considered seriously. The path leading to the final conclusion has also assisted in satisfying the secondary objectives of the project, outlined in the first chapter.

8.2 Viability of the Rank-Conditioned Morphological Transform as a Technique for Generating Line Control Data

The issue of on-line implementation has received limited attention, not having formed part of the initial objectives of the project. Nevertheless, the morphological approaches outlined in Chapter 5, in their final form, are of a level of algorithmic simplicity such that their conversion to a hardware-only system is a relatively straightforward process. The feature extraction operation requiring only three stages:

(i) Set template height according to the grey-tone at the structure's locus (origin).

(ii) Compare each individual pixel intensity value against the threshold level of the corresponding element to assess template containment.

(iii) Shift the analysis structure and prepare it for the next measurement.

The above hardware implementation would offer a real-time response and could provide continuous line control feedback to regulate the annealing environment. Given the ability to implement the algorithms as a highly parallel, region-based processing method, with scalable granularity, such an approach is highly attractive once proven. Providing the lighting condition remains stable, and the capture apparatus is able to offer the same or improved level of textural detail and clarity as the x500 sample set, then an on-line classification rate of at least 80% can be expected.

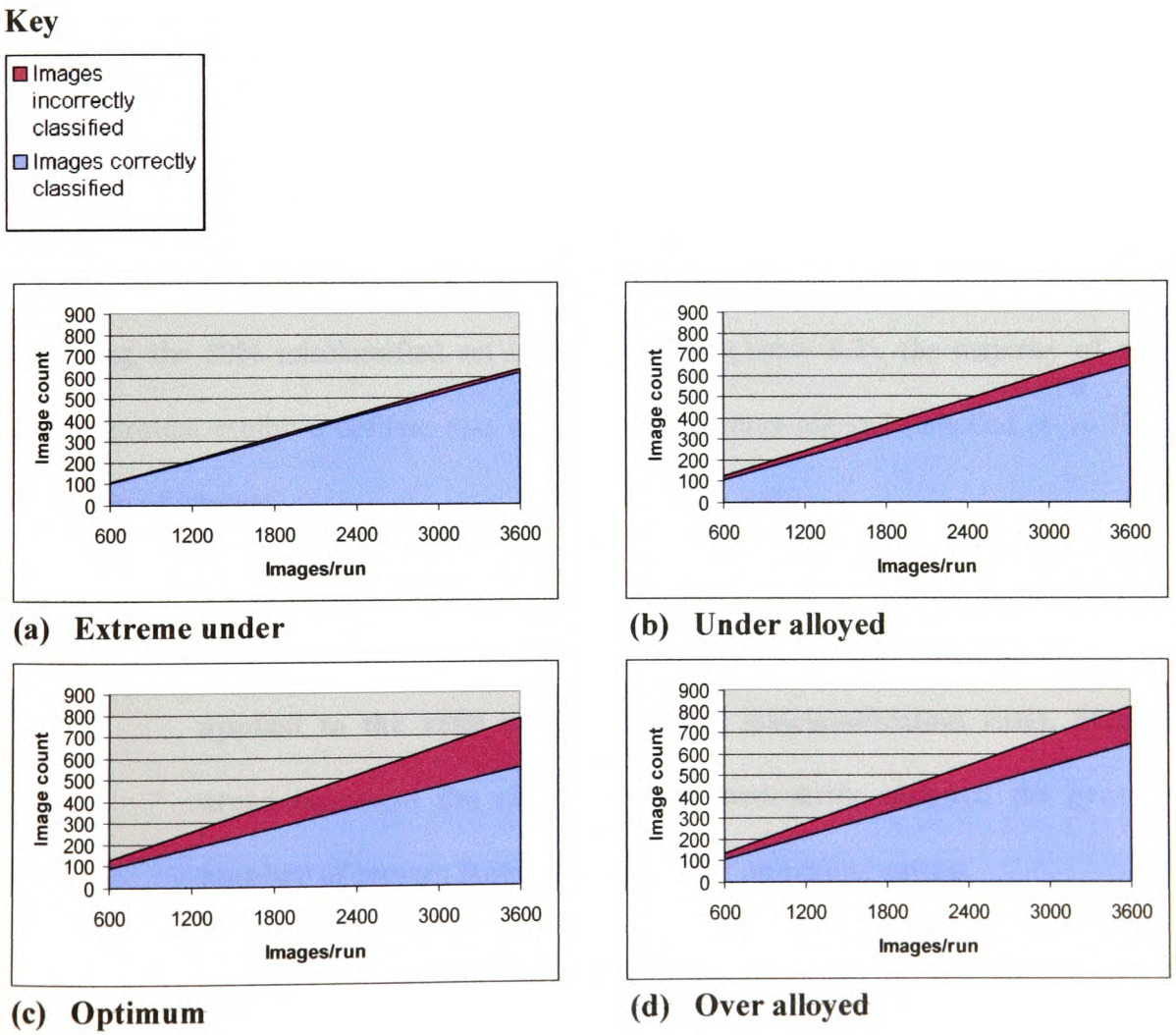
Table 8.1 separates the 80% of the five spiral rank-conditioned morphological result into a series of probabilities that define the expected image assignment precision standard, based on the combination of the number of samples per class and the ratio of correct and incorrect class assignments.

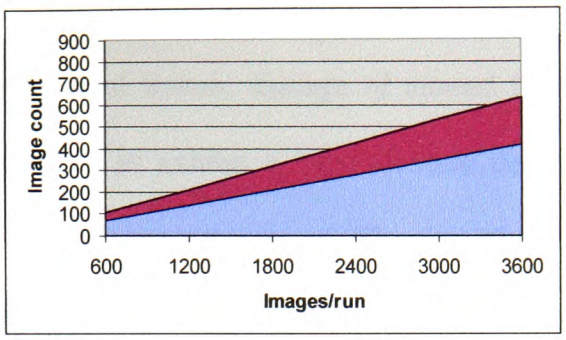
Table 8.1: Summary of classification for the five spiral voting system as applied to the x500 image set (80% classification rate), outlining the probability of classifying each sample on a class-by-class basis

ANNEALING CLASS	YES	NO	IMAGE TOTAL	NUMBER OF IMAGES IN CLASS AS %	PROB. CLASS.	PROB. ERROR
5.38%	68	2	70	17.72%	0.97	0.03
8.42%	71	9	80	20.25%	0.89	0.11
10.79%	60	25	85	21.52%	0.71	0.29
12.08%	71	19	90	22.78%	0.79	0.21
14.37%	46	24	70	17.72%	0.66	0.34
			395	100%		

Using the above probabilities (Table 8.1, columns 6 and 7) in conjunction with the percentage of images allocated to each annealing group (Column 5), it is possible to evaluate the ability of the on-line feature classifier to accurately assign the incoming test samples to the correct class (Figure 8.1).

Figure 8.1: Number of images correctly and incorrectly classified for each class based on the five spiral voting probabilities and a theoretical capture rate of 10 images/second over a time period ranging from 1-6 minutes





(e) Extreme over

Assuming a capture rate of ten images per second over a time period of between one and six minutes, the 20% level of error represents an insignificant fraction of the overall number of images classified. Based on the predefined parameters, the extreme under-alloyed annealing type (Figure 8.1a) would out-perform the alternative groups, with the optimum (Figure 8.1c) and extreme over-alloyed (Figure 8.1e) achieving the lowest rates of correct classification.

Analysing the 20% misclassified set in further detail (Table 8.2), the majority of steel coating groups exhibit a definite bias towards the optimum-alloyed (labelled as 10.79%) collection of images.

Table 8.2: Summary of misclassification for the five spiral voting system as applied to the x500 image set (20% misclassification rate). Shaded areas represent the class that was incorrectly assigned the greatest number of images from each incoming annealing group

		Assigned Annealing class				
		5.38%	**8.42%**	**10.79%**	**12.08%**	**14.37%**
Incoming Annealing Class	**5.38%**		0	1	0	1
	8.42%	1		5	2	1
	10.79%	2	4		6	13
	12.08%	0	5	8		6
	14.37%	1	1	14	8	

The optimum-alloyed class contains the lowest quality samples of the entire data set, producing an average level of texture damage of around 82%. This significant rate of blurring causes the samples to exhibit a lack of class defining textural detail that is replicated within the alternative test groups. The fact that most of the misclassified images favour the optimum class over any other, remains a significant factor for the on-line classifier. The system has the potential to wrongly assign the incoming sample to the optimum type approximately 7% of the time. However, analysing the previous x set of classifications would not only limit any adverse effects that this may cause, but would also increase the rate of correct assignments by analysing and acting upon any significant annealing trends.

It is reasonable to assume that a measured state long run average of 80% at the correct level permits adequate confidence in the anneal state. Once full control over the image acquisition system has been established, the value of marking apparent misclassified areas is likely to be high (surface homogeneity in the underlying steel strip is not guaranteed), particularly where a series of sequential measurements fall awry from the steady state condition of the strip. The added value of a quality report to the customer cannot be ignored.

8.3 Further Issues to Address

There are a number of further additional points that require detailed consideration prior to the implementation of an on-line classifier.

Firstly, the galvanneal steel coil moves at approximately 1.5 m/s [14]. At this rate of movement coupled with the magnification levels required to produce a suitable digital

representative for classification (at least x200), the process of capturing clear images that are unaffected by any depth of field blur becomes an extremely difficult task. The problem stems from relying on any form of optical system that requires an electrical motor to adjust the analysis lens to sharpen the focus. By the time the lens adjusts to compensate for any perceived activity, the coil would have already travelled some distance in both the horizontal and vertical planes. The jolting movement of the strip as it passes the lens is further exaggerated as the magnification level increases. A key factor that would continue to hamper the capturing process even if the apparatus was to be placed at a high tensile point where vertical movement of the steel strip is minimised, e.g. the accumulator near the exit point (Chapter 1, section 1.2, Figure 1.1), although strobe lighting may help to alleviate the problem.

Secondly, the harsh industrial environment surrounding the galvannealing line promotes the rapid build up of dust and fine particles. Exposing the lens to such a dirty climate would allow the particles to collect over a short period of time and further reduce the quality of the captured images. Placing the optical system within a clear container may offer some protection for the delicate equipment, but would only provide an alternative surface for the dust to collect. Blowing fine jets of air over the receptacle's exterior may reduce the rate of build up, but at some stage a certain degree of human intervention would be required.

The final point that requires attention is the need to assess all the process variables involved in the galvannealing process, and the line control data required to manage them. It is recommended that the response time between a change in any of the variables and the direct effect on the annealing phase of the galvanneal surface be evaluated. It is unlikely

that a real-time response is required since the effect of variable adjustment is non-critical i.e. not immediate [12]. Nevertheless, a detailed review into the significant connections between the environmental changes and annealing phase is essential [7, 13, 15, 150].

A real-time process variable feedback system employed to continually monitor the annealing environment would support the decision made by the morphological routine in its current form and assist in verifying/refining any line control changes made, based on the classification report. Performing environment and classification trend analysis by cross-referencing the process variables with the morphological results may allow the annealing operation to be further refined as the environment-annealing links become stronger and more reliable.

8.4 Conclusion

A number of image processing methods have been approached in an attempt to classify the annealing state of galvanneal steel samples. Only the rank-conditioned morphological transform combined with the grey level co-occurrence image filtration scheme has demonstrated any kind of potential. The 100% and 80% classification rates for the x200 and x500 images respectively demonstrate a confidence in the feature's ability to represent the original steel substrate, despite the limited texture quality. The problem of depth of field blur remains a significant obstacle, especially if it remains part of an on-line classification system. However, the problem has been approached with a novel, elementary image filtering technique that exploits, rather than ignores, the affected regions. This has improved the rate of classification over the standard, unfiltered set but represents an operation that should not have to be relied upon. If the problem of image

quality control can be suitably addressed, then the rate of correct phase assignment can only improve.

The implementation of an on-line galvanneal steel classifier that possesses the ability to automatically control the annealing environment, and ensure that the outgoing steel substrate is always at the optimum state, remains an attainable goal providing the non-trivial factors outlined above can be addressed in a suitable manner.

REFERENCES

[1] I. Daubechies, "Ten Lectures on Wavelets", Capital City Press, Philadelphia, ISBN 0-89871-274-2, 1992.

[2] R. M Haralick, L. G. Shapiro, "Computer and Robot Vision, Volume I", Addison-Wesley, Reading, Massachusetts, ISBN 0-201-10877-1, 1992.

[3] R. M. Haralick, K. S. Shanmugam, I. Dinstein, "Textural Features for Image Classification", IEEE Transaction on Systems, Man, and Cybernetics, vol. SMC-3, no. 6, pp. 610-621, Nov. 1973.

[4] D. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley, 1989.

[5] P, A. Devijver, J. Kittler, "Pattern Recognition – A Statistical Approach", Prentice Hall International, Englewood Cliffs, New Jersey, 1982.

[6] S. J. Rees, "Feature Extraction and Object Recognition using Conditional Morphological Operators", Ph.D. Thesis, University of Glamorgan, Jun. 1997.

[7] K. Watts, "A Study of Quality Control During Galvannealing", Ph.D. Thesis, University of Cardiff, May 1994.

[8] B. M. Gray, C. Belleau, M. A. D'Amico, R. L. Pyle, J. F. Butler, "Development of Hot-Dip Zn-Fe Alloy Coated Steels for Exposed Automotive Applications", 31st Mechanical Working and Steel Processing Conference Proceedings, pp. 3-15, 1989.

[9] C. S. Lin, M. Meshii, "Effect of Steel Chemistry on the Microstructure and Mechanical Properties of the Commercial Galvanneal Coatings", Galvatech '95 Conference Proceedings, California, pp. 477-484, 1995.

[10] C. R. Shastry, S. G. Fountoulakis, "Laboratory Evaluation of Formability, Weldability and Painted Corrosion Performance of Prephosphated Electrogalvanized Sheet Steel", Galvatech '95 Conference Proceedings, California, pp. 627-633, 1995.

[11] "Detection of the Optimum Galvanneal Condition", project notes, British Steel, unpublished.

[12] M. A. Haughton, "The Effect of Aluminium and Iron on the Structure of Galvanized Coatings", Proceedings of the Second International Conference on Hot-Dip Galvanizing, pp. 52-87, 1952.

[13] G. Bilello, N. Azzerri, P. F. Berardi, S. Ramundo, "The New ILP Galvanizing Line in Novi Ligure Works", Galvatech '95 Conference Proceedings, California, pp. 21-26, 1995.

[14] Private communication with V. J. Worner, British Steel.

[15] A. Stadlbauer, F. Rubenzucker, K, Zeman, E. Fuhrmann, "A New Galvannealing Process Controller", Galvatech '95 Conference Proceedings, California, pp. 81-85, 1995.

[16] D. C. Cook, R. S. Tuszynski, H. E. Townsend, "Zinc- Iron Phases Formed on Galvanneal Steel", Hyperfine Interactions vol. 54, pp. 781-786, 1990.

[17] W. van Koesveld, M. Lamberigts, A. van der Heiden, L. Bordignon, "Coating Microstructure Assessment and Control for Advanced Product Properties of Galvannealed IF Steels", Galvatech '95 Conference Proceedings, California, pp. 343-355, 1995.

[18] R. G. Grant, D. C. Cook, "Iron-Zinc Phase Identification in Commercial Galvanneal Steel", Galvatech '95 Conference Proceedings, California, pp. 509-514, 1995.

[19] D. C. Cook, R. G. Grant, "Identification of Multiple Iron Sites in Fe-Zn Binary Alloys", Galvatech '95 Conference Proceedings, California, pp. 497-508, 1995.

[20] H. H. Lee, D. Hiam, "Corrosion Resistance of Galvannealed Steel", Corrosion, vol. 45, no. 10, pp. 852-856, Oct. 1989.

[21] J. M. Mataigne, P. Drillet, J. M. Prat, D. Mareuse, P. Terreaux, M. Guttmann, "Optimized Galvannealed Coating Microstructure for Automotive Applications", Galvatech '95 Conference Proceedings, California, pp. 589-597, 1995.

[22] M. Arimura, M. Urai, J. Iwaya, M. Iwai, "Effects of Press-Forming Factors and Flash Plating on Coating Exfoliation of Galvannealed Steel Sheets", Galvatech '95 Conference Proceedings, California, pp. 733-738, 1995.

[23] V. Leroy, "New Trends in the Development of ZnX and Zn + X Coated Steels with Improved Corrosion Resistance: Theory and Practice", Galvatech '95 Conference Proceedings, California, pp. 279-287, 1995.

[24] K. Hashiguchi, A. Yasuda, T. Hanazawa, M. Ohori, T. Ichida, "Galvanneal Sheet Steel with Excellent Press Formability", ATS Steelmaking Conference, pp. 277-283, 1989.

[25] J. Badger, R. L. Pyle, "LTV Indiana Harbor Works Makes use of The is-2000 Automatic Surface Inspection System for Improved Quality Control of Automotive Galvanneal", Galvatech '95 Conference Proceedings, California, pp. 57-63, 1995.

[26] F. van der Heijden, "Image Based Measurement Systems: Object Recognition and Parameter Estimation", John Wiley and Sons, Chichester, West Sussex, lSBN 0-471-950629, 1995.

[27] R. M Haralick, L. G. Shapiro, "Computer and Robot Vision, Volume II", Addison-Wesley, Reading, Massachusetts, ISBN 0-201-56943-4, 1993.

[28] I. Pitas, "Digital Image Processing Algorithms", Prentice Hall, New York, ISBN 0-13-145814-0, 1993.

[29] R. C. Gonzalez, R. E. Woods, "Digital Image Processing", Addison-Wesley, Reading, Massachusetts, 1992.

[30] D. Patel, T. J. Stonham, "Unsupervised Texture Segmentation and its Application to Real-World Data", Visual Communications and Image Processing, vol. 1818, pp. 1206-1217, 1992.

[31] S. W. Zucker, K. Kant, "Multiple-Level Representations for Texture Discrimination", Pattern Recognition and Image Processing Conference, Chicago, Illinois, pp. 609-614, 1981.

[32] N. Ahuja, A. Rosenfeld, "Mosaic Models for Texture", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 3, no. 1, pp. 1-11, Jan. 1981.

[33] A. C. Bovik, M. Clark, W. S. Geisler, "Multichannel Texture Analysis using Localized Spatial Filters", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp. 55-73, Jan. 1990.

[34] J. W. Modestino, R. W. Fries, A. L. Vickers, "Texture Discrimination Based upon an Assumed Stochastic Texture Model", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 3, no. 5, pp. 557-580, Sep. 1981.

[35] R. M. Picket, "Visual Analysis of Texture in the Detection and Recognition of Objects", in Picture Processing and Psychopictorics, Editors: B. S. Lipkin, A. Rosenfeld, Academic Press, New York, pp. 289-308, 1970.

[36] J. K. Hawkins, "Textural Properties for Pattern Recognition", in Picture Processing and Psychopictorics, Editors: B. S. Lipkin, A. Rosenfeld, Academic Press, New York, pp. 347-370, 1970.

[37] W. K. Pratt, "Digital Image Processing", John Wiley and Sons, New York, 1978.

[38] R. M. Haralick, "A Resolution Preserving Textural Transform for Images", Proceedings of the IEEE Conference on Computer Graphics, Pattern Recognition and Data Structures, Beverly Hills, California, pp. 51-61, 1975.

[39] M. Dentith, "Textural Filtering of Aeromagnetic Data", Exploration Geophysics, vol. 26, pp. 209-214, 1995.

[40] A. Rosenfeld, E. Troy, "Visual Texture Analysis", Conference Record for Symposium on Feature Extraction and Selection in Pattern Recognition Visual Texture Analysis, Argonne, Illinois, pp 115-124, 1970.

[41] K. Wiltschi, T. Lindeberg, A. Pinz, "Classification of carbide Distributions using Scale Selection and Directional Distributions", Technical Report ISRN KTH/NA/P-97/10-SE, Oct. 1997.

[42] K. Wiltschi, A. Pinz, T. Lindeberg, "An Automatic Assessment Scheme for Steel Quality Inspection", Technical Report ISRN KTH/NA/P-98/20-SE, Nov. 1998.

[43] A. K. Louis, P. Maaß, A. Rieder, "Wavelets: Theory and Applications", John Wiley and Sons, Chichester, West Sussex, ISBN 0-471-96792-0, 1997.

[44] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, "Numerical Recipes in Fortran", Cambridge University Press, Cambridge, ISBN 0-521-43064-X, 1992.

[45] Y. Nievergelt, "Wavelets Made Easy", Birkhäuser, Boston, 1999.

[46] R. A. DeVore, B. Jawerth, B. J. Lucier, "Image Compression Through Wavelet Transform Coding", IEEE Transactions on Information Theory, vol. 38, no. 2, Mar. 1992.

[47] R. Öktem, K. Egiazarian, "A Wavelet Transform Method for Coding Film-Grain Noise Corrupted Images", IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, Jun. 2000.

[48] R. Öktem, K. Egiazarian, "A Transform Domain Method for Filtering Film-Grain Type Noise", Proceedings of the Second International Conference on Information, Communications and Signal Processing, Singapore, Dec. 1999.

[49] R. Öktem, K. Egiazarian, "Transform Domain Algorithm for Reducing the Effect of Film-Grain Noise in Image Compression", Electronics Letters, vol.35, no. 21, pp. 1830-1831, 1995.

[50] J. L. Starck, F. Murtagh, "Image Restoration with Noise Suppression using the Wavelet Transform", Astronomy and Astrophysics, vol. 288, pp. 342-348, 1994.

[51] J. L. Starck, F. Murtagh, A. Bijaoui, "Image Processing and Data Analysis – The Multiscale Approach", Cambridge University Press, Cambridge, ISBN 0-521-59084-1, 1998.

[52] S. Theodoridis, K. Koutroumbas, "Pattern Recognition", Academic Press, 1999.

[53] R. J. Schalkoff, "Digital Image Processing and Computer Vision", John Wiley and Sons, New York, 1989.

[54] S. T. Bow, "Pattern Recognition: Applications to Large Data-Set Problems", Marcel Dekker, New York, 1984.

[55] A. R. Weeks Jr, "Fundamentals of Electronic Image Processing", SPIE Optical Engineering Press, Bellingham, Washington, ISBN 0-8194-2149-9, 1996.

[56] M. L. Malagnini, M. Pucillo, P. Santin, "A Project for Faint Object Discrimination in Astronomy", in S. Levialdi, editor, "Digital Image Processing", pp. 343-354, Pitman Publishing, 1991.

[57] E. R. Dougherty, C. R. Giardina, "Matrix Structured Image Processing", Prentice Hall, Englewood Cliffs, New Jersey, ISBN 0-13-565623-0, 1987.

[58] R. M. Haralick, "Statistical and Structural Approaches to Texture", Proceedings of the IEEE, vol. 67, no. 5, pp. 786-804, May 1979.

[59] R. M. Haralick, K. S. Shanmugam, "Computer Classification of Reservoir Sandstones", IEEE Transactions on Geoscience Electronics, vol. 11, pp. 171-177, 1973.

[60] R. M. Haralick, "On a Texture-Context Feature Extraction Algorithm for Remotely Sensed Imagery", IEEE Decision and Control Conference, Miami, Florida, pp. 650-657, 1971.

[61] R. M. Haralick, K. S. Shanmugam, "Combined Spectral and Spatial Processing of ERTS Imagery Data", Journal of Remote Sensing of the Environment, vol. 3, pp. 3-13, 1974.

[62] D. Terzopoulos, S. W. Zucker, "Detection of Osteogenesis Imperfecta by Automated Texture Analysis", Computer Graphics and Image Processing, vol. 20, pp. 229-243, 1982.

[63] R. F. Walker, P. Jackway, B. Lovell, "Cervical Cell Classification via Co-Occurrence and Markov Random Field Features", Proceedings of Digital Image Computing: Techniques and Applications, pp. 294-299, 1995.

[64] G. S. Cox, F. J. Hoare, G. de Jager, "Experiments in Lung Cancer Nodule Detection using Texture Analysis and Neural Network Classifiers", Third South African Workshop on Pattern Recognition, Pretoria, 1992.

[65] L. H. Siew, R. M. Hodgson, E. J. Wood, "Texture Measures for Carpet Wear Assessment", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 1, pp. 92-105, Jan. 1988.

[66] R. W. Conners, C. A. Harlow, "Some Theoretical Considerations Concerning Texture Analysis of Radiographic Images", Proceedings of the 1976 Conference on Decision and Control, Clearwater, Florida, 1976.

[67] L. Van Gool, P. Dewaele, A. Oosterlinck, "Texture Analysis Anno 1983", Computer Vision, Graphics, and Image Processing, vol. 29, pp. 336-357, 1985.

[68] J. H. Holland, "Adaptation in Natural and Artificial Systems", University of Michigan Press, Michigan, 1975.

[69] L. Davis, "The Handbook of Genetic Algorithms", Van Nostrand Reingold, New York, ISBN 0-442-00173-8, 1991.

[70] S. A. Dudani, "The Distance-Weighted K-Nearest Neighbour Rule", IEEE Transaction on Systems, Man, and Cybernetics, pp. 325-327, Apr. 1976.

[71] W. F. Punch, E. D. Goodman, M. Pei, L. C. Shun, P. Hovland, R. Enbody, "Further Research on Feature Selection and Classification using Genetic Algorithms", Proceedings of the Fifth International Conference on Genetic Algorithms, San Mateo, California, pp. 557-564, 1993.

[72] W. Siedlecki, J. Sklansky, "A Note on Genetic Algorithms for Large-Scale Feature Selection", Pattern Recognition Letters, vol. 10, pp. 335-347, 1989.

[73] M. L. Raymer, P. C. Sanschargrin, W. F. Punch, S. Venkataraman, E. D. Goodman, L. A. Kuhn, "Predicting Conserved Water-Mediated and Polar Ligand Interactions in Proteins using a K-nearest Genetic Algorithm", Journal of Molecular Biology, vol. 265, pp. 445-464, 1997.

[74] J. D. Kelly Jr, L. Davis, "Hybridizing the Genetic Algorithm and the K Nearest Neighbours Classification Algorithm", Proceedings of the Fourth International Conference on Genetic Algorithms and their Applications, San Diego, California, pp. 377-383, 1991.

[75] J. D. Kelly Jr, L. Davis, "A Hybrid Genetic Algorithm for Classification", Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, Sydney, pp. 645-650, Aug. 1991.

[76] S. J. Rees, B. F. Jones, T. D. Jones, "Direct Feature Extraction using Conditional Morphological Operators", Proceedings of the Tenth International Conference on Systems Engineering, Coventry, pp. 998-1005, Sep. 1994.

[77] S. J. Rees, B. F. Jones, "Conditional Morphological Operators for Direct Feature Extraction and Enhancement", Proceedings of the IEE Fifth International Conference on Image Processing and its Applications, IEE Conference Publication No. 410, Edinburgh, pp. 747-751, Jul. 1995.

[78] S. J. Rees, S. L. Woodham, L. S. Dooley, "Use of Rank-Conditioned Morphological Operators for Texture Classification", Proceedings of the IEE Sixth International Conference on Image Processing and its Applications, Dublin, pp. 618-622, Jul. 1997.

[79] P. Soille, "Morphological Image Analysis: Principles and Applications", Springer-Verlag, Berlin Heidelberg, ISBN 3-540-65671-5, 1999.

[80] P. Salembier, "Structuring Element Adaptation for Morphological Filters", Journal of Visual Communication and Image Representation, vol. 3, no. 2, pp. 115-136, Jun. 1992.

[81] R. P. Loce, E. R. Dougherty, "Facilitation of Optimal Binary Morphological Filter Design via Structuring Element Libraries and Design Constraints", Optical Engineering, vol. 31, no. 5, pp. 1008-1025, May 1992.

[82] E. R. Dougherty, R. P. Loce, "Optimal Mean-Absolute-Error Hit-or-Miss Filters: Morphological Representation and Estimation of the Binary Conditional Expectation", Optical Engineering, vol. 32, no. 4, pp. 815-827, Apr. 1993.

[83] R. Ehrhardt, "Morphological Filter Design with Genetic Algorithms", SPIE, vol. 2300, pp. 2-12, 1994.

[84] N. R. Harvey, S. Marshall, "The Design of Different Classes of Morphological Filter using Genetic Algorithms", Proceedings of the IEE Fifth International Conference on Image Processing and its Applications, IEE Conference Publication No. 410, Edinburgh, pp. 227-231, Jul. 1995.

[85] E. R. Dougherty, J. T. Astola, "Nonlinear Filters for Image Processing", SPIE, IEEE, 1999.

[86] W. Li, V. Haese-Coat, "Composite Morphological Filters in Multiresolution Morphological Decomposition", Image Processing and its Applications, no. 410, Jul. 1995.

[87] R. L. Kashyap, A. Khotanzad, "A Model-Based Method for Rotation Invariant Texture Classification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 4, pp. 472-481, Jul. 1986.

[88] S. W. Zucker, "Finding Structure in Co-Occurrence Matrices for Texture Analysis", Computer Graphics and Image Processing, vol. 12, pp. 286-308, 1980.

[89] N. Temme, "An Introduction to Wavelets", CWI, Nov. 1997.

[90] S. G. Sánchez, N. G. Prelcic, S. J. G. Galán, Uvi-Wave Toolbox Manual, 1996.

[91] B. Vidakovic, P. Muller, "Wavelets for Kids", ftp://ftp.isds.duke.edu/pub/brani/papers/wav4kids[A-B].ps.z, 1994.

[92] "Wavelets and their Applications", Siggraph '95 course notes, 1995.

[93] W. Sweldens, "The Construction and Application of Wavelets in Numerical Analysis", Ph.D. Thesis, May 1995.

[94] P. J. Oonincx, "Automatic Phase Detection in Seismic Data using the Discrete Wavelet Transform", Probability, Networks and Algorithms, PNA-R 9811, Oct. 1998.

[95] J. M. Lina, L. Gagnon, "Image Enhancement with Symmetric Daubechies Wavelets", Nuclear Physics Lab., University of Montréal, 1995.

[96] J. Lu, "Signal Recovery and Noise Reduction using Wavelets", Ph.D. Thesis, Dartmouth College, Jun. 1993.

[97] J. M. Lina, "Image Processing with Complex Daubechies Wavelets", Centre for Mathematical Research, University of Montréal, 1995.

[98] J. P. Bolet, A. R. Cowen, J. Launders, A. G. Davies, G. J. S. Parkin, R. F. Bury, "Progress with an "All-Wavelet" Approach to Image Enhancement and De-Noising of Direct Digital Thorax Radiographic Images", Proceedings of the IEE Sixth International Conference on Image Processing and its Applications, Dublin, pp. 244-248, Jul. 1997.

[99] S. Pittner, S. Kamarthi, "Feature Extraction from Wavelet Coefficients for Pattern Recognition Tasks", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 1, pp. 83-88, Jan. 1999.

[100] J. B. T. M. Roerdink, M. A. Westenberg, "Wavelet-Based Volume Visualization", Proceedings Dutch Mathematical Congress, 1998.

[101] J. L. Starck, F. Murtagh, "Multiresolution Image Analysis using Wavelets – Recent Results", Bulletin of the American Astronomical Society, vol. 26, pp. 1003-1005, 1994

[102] F. C. Cesbron, F. J. Malassenet, "Wavelet and Fractal Transforms for Image Compression", Proceedings of the IEE Sixth International Conference on Image Processing and its Applications, Dublin, pp. 77-80, Jul. 1997.

[103] M. L. Hilton, B. D. Jawerth, A. Sengupta, "Compressing Still and Moving Images with Wavelets", Multimedia Systems, vol. 2, no. 3, Apr. 1994.

[104] A. Grapps, "An Introduction to Wavelets", IEEE Computational Science and Engineering, vol. 2, no. 2, 1995.

[105] P. M. Bentley, J. T. E. McDonnell, "Wavelet Transforms: An Introduction", Electronics and Communication Engineering Journal, pp. 175-186, Aug. 1994.

[106] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets", Communications on Pure and Applied Mathematics, vol. 41, no. 7, pp. 909-996, 1988.

[107] M. A. Cody, "The Wavelet Packet Transform", Dr. Dobb's Journal, pp. 44-54, Apr. 1994.

[108] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, "Image Coding using Wavelet Transform", IEEE Transactions on Image Processing, vol. 1, no. 2, pp. 205-220, Apr. 1992.

[109] R. R. Coifman, M. V. Wickerhauser, "Entropy-Based Algorithms for Best Bases Selection", IEEE Transaction on Information Theory, vol. 38, no. 2, pp. 713-718, Mar. 1992.

[110] J. H. P. Burrill, "Texture Mapping of Neurological Magnetic Resonance Images", Individual Project, Department of Electric and Electronic Engineering, Imperial College of Science, London, 1997.

[111] R. Finn, "Quantitve Assessment of Neurological Image Characteristics in Alzheimer's Patients", MSc. Thesis, Imperial College Centre for Biological and Medical Systems, 1993.

[112] A. R. Al-hinnawi, P. E. Windrill, G. Needham, "The Use of Image Texture Analysis in the Detection of Micro-Calculations within the Mammogram", Proceedings of the IEE Sixth International Conference on Image Processing and its Applications, Dublin, pp. 283-288, Jul. 1997.

[113] M. D. Levine, "Vision in Man and Machine", McGraw-Hill, New York, ISBN 0-07-037446-5, 1985.

[114] P. P. Ohanian, R. C. Dubes, "Performance Evaluation for Four Classes of Textured Features", Pattern Recognition, vol. 25, no. 8, pp 819-833, 1992.

[115] R. J. Tully, R. W. Conners, C. A. Harlow, G. S. Lodwick, "Towards Computer Analysis of Pulmonary Infiltration", Investigative Radiology, vol. 13, pp. 298-305, 1978.

[116] R. W. Conners, C. A. Harlow, "A Theoretical Comparison of Texture Algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 2, no. 3, pp. 204-222, May 1980.

[117] R. F. Walker, P. Jackway, I. D. Longstaff, "Improving Co-occurrence Matrix Feature Discrimination", Proceedings of the Third Conference on Digital Image Computing: Techniques and Applications, pp 643-648, Dec. 1995.

[118] G. Smith, "Variations of GLCM – Varying Distance Parameter", http://www.cssip.elec.uq.edu.au/~guy/meastex/www/comparisons/glcmDistance.html, guy@it.uq.edu.au, 1998.

[119] R. Schalkoff, "Pattern Recognition: Statistical, Structural and Neural Approaches", John Wiley and Sons, New York, ISBN 0-471-52974-5, 1992.

[120] Jasc Software, Paint Shop Pro Version 5.00, 1998.

[121] C. Watkins, A. Sadun, S. Marenka, "Modern Image Processing: Warping, Morphing, and Classical Techniques", Academic Press, Cambridge, Massachusetts, ISBN 0-12-737860-X, 1993.

[122] J. Grefenstette, R. Gopal, B. Rosmaita, D. V. Gucht, "Genetic Algorithms for the Travelling Salesman Problem", Proceedings of the First International Conference on Genetic Algorithms and their Applications, Hillsdale, New Jersey, pp. 160-168, 1985.

[123] W, Cedeño, V. R. Vemuri, "Genetic Algorithms in Aquifer Management", Journal of Network and Computer Applications, vol. 19, pp. 171-187, 1996.

[124] I, Ashiru, C. Czarnecki, T. Routen, "Characteristics of a Genetic Based Approach to Path Planning for Mobile Robots", Journal of Network and Computer Applications, vol. 19, pp. 149-169, 1996.

[125] M. Morrow, "Genetic Algorithms: A New Class of Searching Algorithms", Dr. Dobb's Journal, pp. 26-32, Apr. 1991.

[126] K. A. DeJong, "Analysis of the Behaviour of a Class of Genetic Adaptive Systems", Ph.D. Thesis, University of Michigan, 1975.

[127] R. B. Hollstien, "Artificial Genetic Adaptation in Computer Controlled Systems", Ph.D. Thesis, University of Michigan, 1971.

[128] D. Beasley, D. R. Bull, R. R. Martin, "An Overview of Genetic Algorithms: Part 1, Fundamentals", University Computing, vol. 15, no. 2, pp. 58-69, 1993.

[129] A. Hunter, "SUGAL User Manual v2.1", http://osiris.sunderland.ac.uk/ahu/sugal/home.html, cs0ahu@sunderland.ac.uk, Jul. 1995.

[130] Z. Michalewicz, "Genetic Algorithms + Data Structures = Evolution Programs", Springer-Verlag, New York, 1994.

[131] F. Hoffmeister, T. Bäck, "Genetic Algorithms and Evolution Strategies: Similarities and Differences", Proceedings of the First International Conference on Parallel Problem Solving from Nature, vol. 496, pp. 455-469, 1992.

[132] C. Z. Janikow, Z. Michalewicz, "An Experimental Comparison of Binary and Floating Point Representations in Genetic Algorithms", Proceedings of the Fourth International Conference on Genetic Algorithms and their Applications, San Diego, California, pp. 31-36, 1991.

[133] A. H. Wright, "Genetic Algorithms for Real Parameter Optimization", in G. J. E. Rawlins, editor, "Foundations of Genetic Algorithms", pp. 205-218, Morgan Kaufmann Publishers, 1991.

[134] J. Antonisse, "A new Interpretation of Schema Notion that Overturns the Binary Encoding Constraint", Proceedings of the Third International Conference on Genetic Algorithms, Los Altos, California, pp. 86-91, 1989.

[135] D. Greenhalgh, S. Marshall, "An Improved Stopping Criteria for Genetic Algorithms", Sicom Journal of Computing, Jun. 2000.

[136] J. Heitkoetter, D. Beasley, "Frequently Asked Questions for Evolutionary Computation", http://www.faqs.org/faqs/ai-faq/genetic/, (parts 1 to 6 available), david.beasley@iee.org, 1997.

[137] D. Beasley, D. R. Bull, R. R. Martin, "An Overview of Genetic Algorithms: Part 2, Research Topics", University Computing, vol. 15, no. 4, pp. 170-181, 1993.

[138] L. J. Eshelman, R. Caruna, J. D. Schaffer, "Biases in the Crossover Landscape", Proceedings of the Third International Conference on Genetic Algorithms and their Applications, San Mateo, California, pp. 10-19, 1989.

[139] R. O. Duda, P. E. Hart, "Pattern Classification and Scene Analysis", John Wiley and Sons, New York, 1973.

[140] M, L. Raymer, P. C. Sanschagrin, W. F. Punch, S. Venkataraman, E. D. Goodman, L. A. Kuhn, "Predicting Conserved Water-mediated and Polar Ligand Interactions in Proteins Using a K-nearest-neighbors Algorithm", Journal of Molecular Biology, vol. 265, pp. 445-464, 1997.

[141] S. J. Raudys, A. K. Jain, "Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 3, pp. 252-264, 1991.

[142] D. R. Anderson, D. J. Sweeney, T. A. Williams, "Introduction to Statistics – Concepts and Applications", West Publishing Company, St. Paul, Minneapolis, ISBN 0-314-02813-7, 1991.

[143] D. A. Berry, B. W. Lindgren, "Statistics – Theory and Methods", Duxbury Press, Belmont, California, ISBN 0-534-50479-5, 1996.

[144] S. R. Sternberg, "Grayscale Morphology", Computer Vision, Graphics, and Image Processing, no. 35, pp. 333-355, April 1986.

[145] R. M. Haralick, S. R. Sternberg, X. Zhuang, "Image Analysis using Mathematical Morphology", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 9, no. 4, pp. 532-549, July 1987.

[146] E. R. Dougherty, R. M. Haralick, "The Hole Spectrum - Model-Based Optimization of Morphological Filters", Image Algebra and Morphological Processing II, vol. 1568, pp. 224-232, 1991.

[147] E. R. Dougherty, A. Mathew, V. Swarnakar, "A Conditional-Expectation-Based Implementation of the Optimal Mean-Square Binary Morphological Filter", Non-linear Image Processing II, vol. 1451, pp. 137-147, 1991.

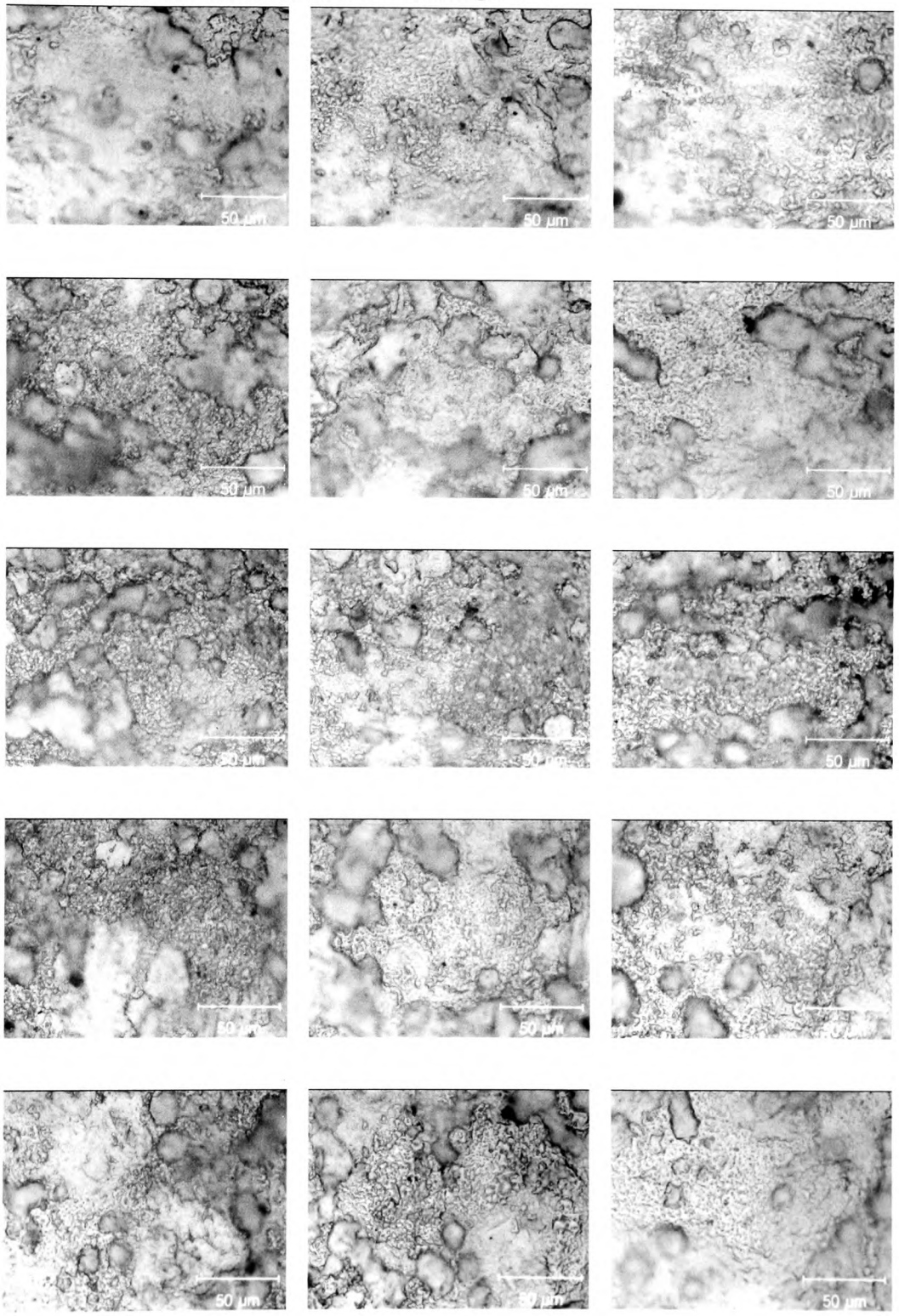
[148] P. Brodatz, "Textures – A Photographic Album for Artists and Designers", Dover, New York, 1966.

[149] R. W. Conners, M. H. Trivedi, C. A. Harlow, "Segmentation of a High-Resolution Urban Scene Using Texture Operators", Computer Vision, Graphics, and Image Processing, vol. 25, pp. 273-310, 1984.

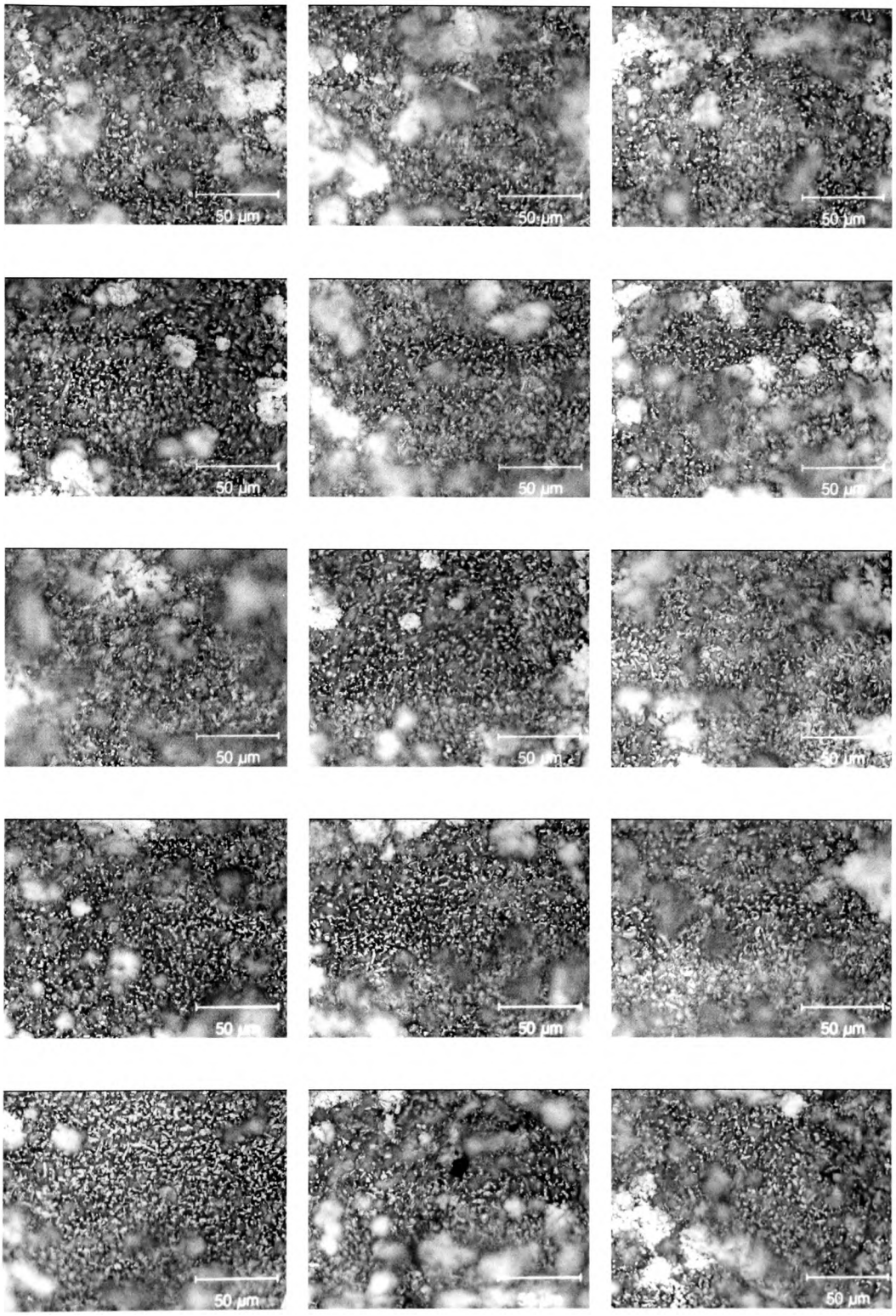
[150] Y. Suemitsu, "New Coating Line Technologies", Galvatech '95 Conference Proceedings, California, pp. 3-12, 1995.

Appendix 1

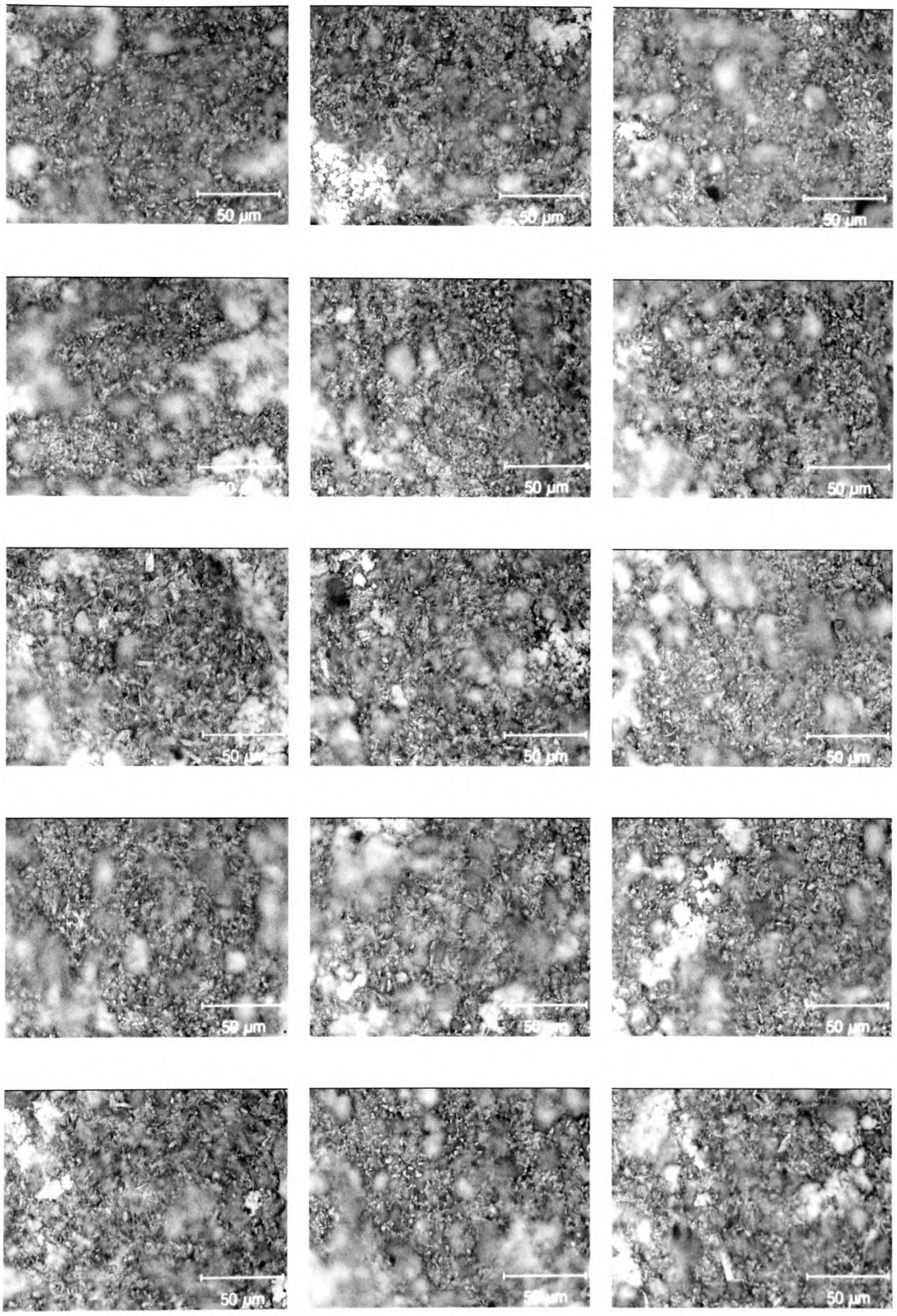
x500 image samples drawn from each annealing class



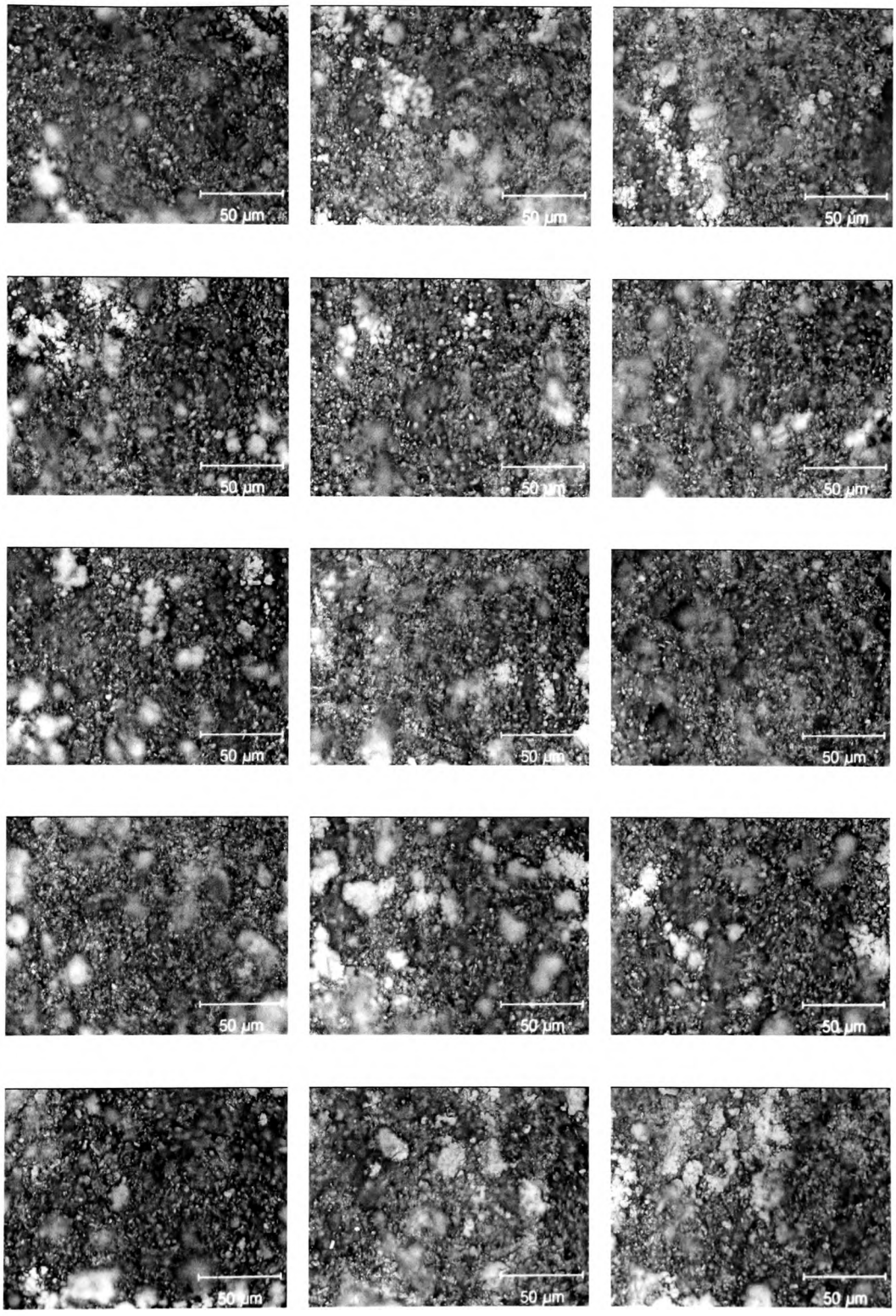
37.6 g/m, 5.38% Fe, 0.66% Al
(Extreme under-alloyed)



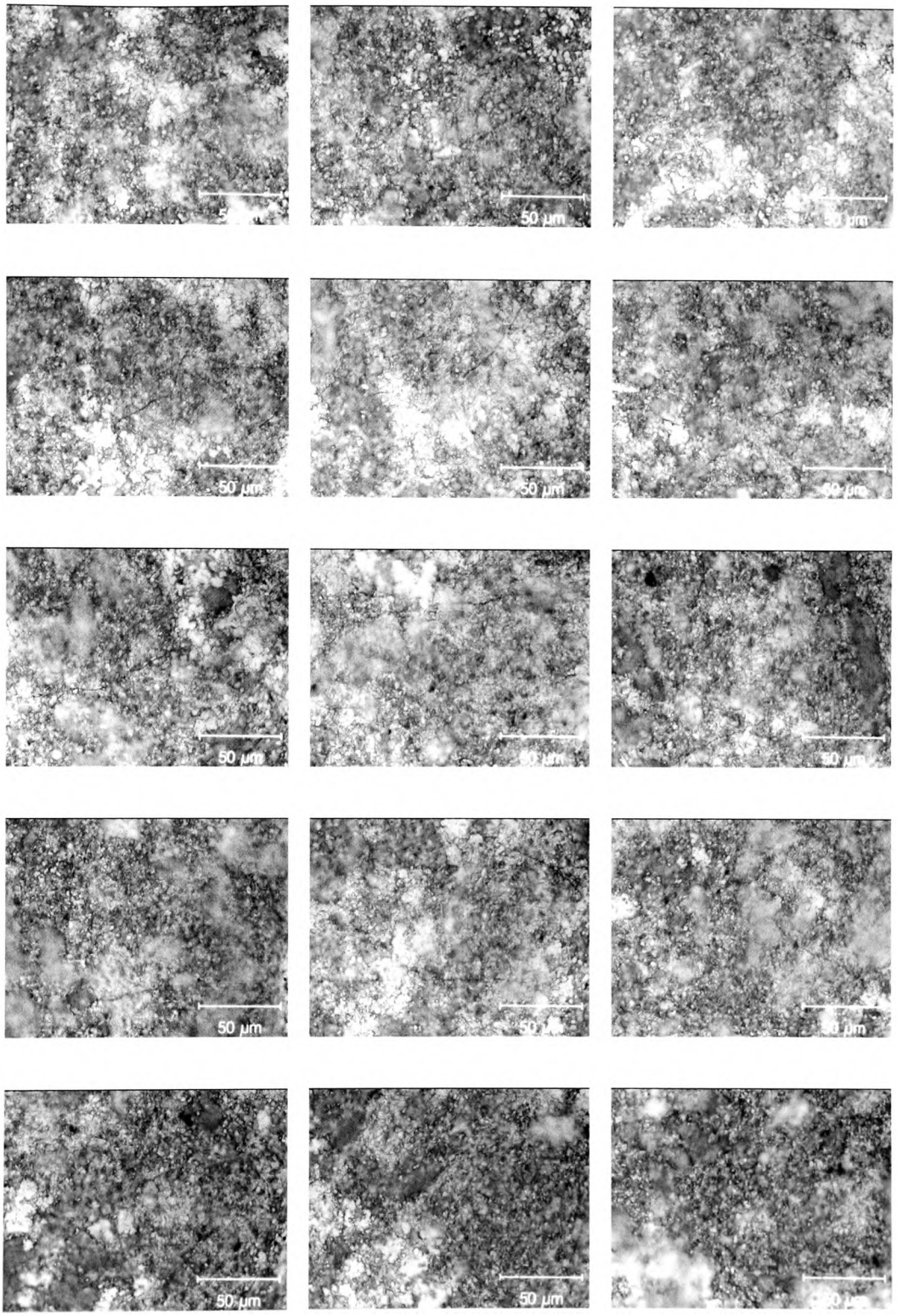
41 g/m, 8.42% Fe, 0.61% Al
(Under-alloyed)



41.9 g/m, 10.79% Fe, 0.59% Al
(Optimum-alloyed condition)



43 g/m, 12.08% Fe, 0.6% Al
(Over-alloyed)



47.5 g/m, 14.37% Fe, 0.54% Al
(Extreme over-alloyed)

Appendix 2

Statistical measurements based on the GLCM descriptors

Probability of assigning the incoming data to the correct template group and the resulting probability of error. Shaded areas represent matching subsets extracted from the same class. For example, using the correlation measurement, test data group A has a probability of being correctly assigned to group B of 0.725 and a probability of 0.275 of being wrongly classified as group J, hence the probability of error for test data A is 0.275.

Correlation

		Training Data					
		B	**D**	**F**	**H**	**J**	**Probability of Error**
Test data	**A**	0.725	0	0	0	0.275	0.275
	C	0	0.3580247	0.345679	0.0987654	0.1975309	0.6419753
	E	0	0.3098592	0.4225352	0	0.2676056	0.5774648
	G	0	0.3111111	0.2555556	0.3111111	0.1222222	0.6888889
	I	0.0649351	0.1948052	0.3636364	0	0.3766234	0.6233766

Energy

		Training Data					
		B	**D**	**F**	**H**	**J**	**Probability of Error**
Test data	**A**	0.2923077	0.1076923	0.2461539	0.3076923	0.0461539	0.7076923
	C	0.2056738	0.2056738	0.2056738	0.212766	0.1702128	0.7943262
	E	0.2168675	0.1445783	0.2891566	0.2771084	0.0722892	0.7108434
	G	0.2047244	0.1889764	0.2283465	0.2362205	0.1417323	0.7637795
	I	0.1428571	0.2222222	0.2142857	0.2063492	0.2142857	0.7857143

Entropy

Test data	Training Data: B	D	F	H	J	Probability of Error
A	0.32	0.0666667	0.2933333	0.2666667	0.0533334	0.68
C	0.1944444	0.1944444	0.2013889	0.2083333	0.2013889	0.8055556
E	0.2168675	0.0843374	0.2891566	0.2771084	0.1325301	0.7108434
G	0.2288136	0.1271186	0.2457627	0.2457627	0.1525424	0.7542373
I	0.1567164	0.2089552	0.2089552	0.2164179	0.2089552	0.7910448

Element Difference of Order 3

Test data	Training Data: B	D	F	H	J	Probability of Error
A	0.4210526	0	0.3157895	0.1754386	0.0877193	0.5789474
C	0.1702128	0.1985816	0.2056738	0.212766	0.212766	0.8014184
E	0.2816902	0.0140845	0.3380282	0.1971831	0.1690141	0.6619718
G	0.1960784	0.0490196	0.2647059	0.2843137	0.2058824	0.7156863
I	0.2222222	0.1111111	0.2222222	0.2222222	0.2222222	0.7777778

Element Difference of Order 4

Test data	Training Data: B	D	F	H	J	Probability of Error
A	0.52	0	0.34	0.04	0.1	0.48
C	0.141791	0.2014925	0.2089552	0.2238806	0.2238806	0.7985075
E	0.25	0.0147588	0.3676471	0.1323529	0.2352941	0.6323529
G	0.1346154	0.0769231	0.2307692	0.2788461	0.2788461	0.7211539
I	0.2230769	0.0846154	0.2307692	0.2307692	0.2307692	0.7692308

Inverse Difference Moment

Test data	Training Data: B	D	F	H	J	Probability of Error
A	0.2537313	0.1492537	0.2835821	0.2686567	0.0447761	0.7462686
C	0.24	0.23	0.23	0.25	0.05	0.77
E	0.25	0.2	0.2625	0.2625	0.05	0.7375
G	0.2093023	0.2248062	0.2248062	0.2325581	0.1085271	0.7674419
I	0.1846154	0.2230769	0.2153846	0.2153846	0.1615385	0.8384615

Inertia

		Training Data					
		B	D	F	H	J	Probability of Error
Test data	A	0.3333333	0.0166667	0.3	0.2833333	0.0666667	0.6666666
	C	0.1931034	0.1931034	0.2	0.2068966	0.2068966	0.8068966
	E	0.2835821	0.0298508	0.3134328	0.2985075	0.0746269	0.6865672
	G	0.2342342	0.0810811	0.2522523	0.2612613	0.1711712	0.7387388
	I	0.2164179	0.119403	0.2164179	0.2238806	0.2238806	0.7761194

Maximum Probability

		Training Data					
		B	D	F	H	J	Probability of Error
Test data	A	0.226087	0.1217391	0.2	0.2086957	0.2434783	0.773913
	C	0.2027972	0.1818182	0.2027972	0.2027972	0.2027972	0.8181818
	E	0.2142857	0.1349206	0.2142857	0.1984127	0.2380952	0.7857143
	G	0.2014925	0.1716418	0.2089552	0.1940299	0.2238806	0.8059701
	I	0.2054795	0.1849315	0.1986301	0.2054795	0.2054795	0.7945206

Prominence

		Training Data					
		B	D	F	H	J	Probability of Error
Test data	A	0.6304348	0.0217391	0.0434783	0	0.3043478	0.3695652
	C	0	0.3095238	0.3452381	0.0952381	0.25	0.6904762
	E	0	0.3139535	0.3488372	0.0813954	0.255814	0.6511628
	G	0	0.2843137	0.2745098	0.2843137	0.1568628	0.7156863
	I	0.0886076	0.2151899	0.3164557	0	0.3797468	0.6202532

Shade

		Training Data					
		B	D	F	H	J	Probability of Error
Test data	A	0.6590909	0.0227273	0.0454546	0	0.2727273	0.3409091
	C	0	0.3375	0.3625	0.0625	0.2375	0.6625
	E	0	0.2948718	0.3846154	0.0512821	0.2692308	0.6153846
	G	0	0.2959184	0.2755102	0.2857143	0.1428571	0.7142857
	I	0.0657895	0.2105263	0.3421053	0	0.381579	0.6184211

Number of mean and variance matches for each GLCM measurement within each class, based on the standard and F distributions respectively. Shaded boxes represent the two subset images from the same class, i.e. A and B, C and D, E and F etc. are of the same type.

Example, at subset 1, reference E and subset 2, reference F, for the mean comparisons using the standard probability distribution, 5 out of the 10 GLCM measurements were considered as possessing the same mean value. Similarly, for the variance comparisons using the F distribution, 3 out of the 10 measurements have matching variances.

Note: These tables do not specify which measurements match.

Frequency count of the number of cluster-mean values that match

		Subset Group 1									
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Subset Group 2	**A**		1	1	1	0	1	0	0	1	1
	B	1		2	1	7	3	3	2	1	1
	C	1	2		5	2	3	2	2	3	2
	D	1	1	5		1	2	2	2	3	2
	E	0	7	2	1		5	2	2	1	1
	F	1	3	3	2	5		5	5	4	1
	G	0	3	2	2	2	5		10	3	2
	H	0	2	2	2	2	5	10		3	1
	I	1	1	3	3	1	4	3	3		4
	J	1	1	2	2	1	1	2	1	4	

Frequency count of the number of cluster-variance values that match.

		Subset Group 1									
		A	B	C	D	E	F	G	H	I	J
Subset Group 2	A		5	1	1	4	5	2	4	2	3
	B	5		3	1	2	5	4	1	4	3
	C	1	3		9	4	4	5	6	6	2
	D	1	1	9		6	4	6	6	6	4
	E	4	2	4	6		3	2	3	4	5
	F	5	5	4	4	3		7	7	4	4
	G	2	4	5	6	2	7		9	3	6
	H	4	1	6	6	3	7	9		4	6
	I	2	4	6	6	4	4	3	4		5
	J	3	3	2	4	5	4	6	6	5	

Appendix 3

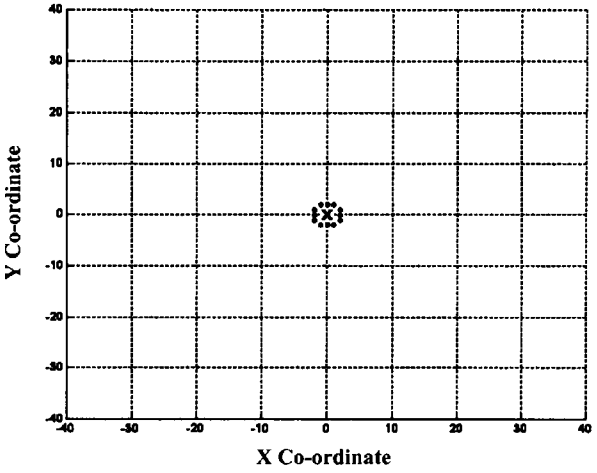
Rank-conditioned morphological template set

Sparse annular collection - locus (origin) represented by centre point x

(co-ordinate and pictorial views)

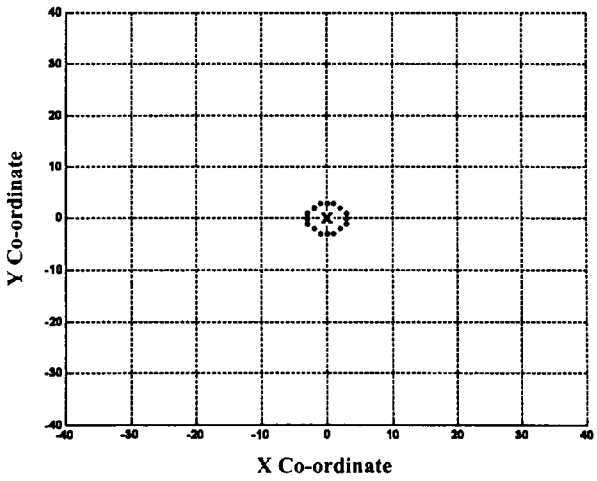
Small nut

X CO-ORDINATE	Y CO-ORDINATE
-1	2
0	2
1	2
2	1
2	0
2	-1
1	-2
0	-2
-1	-2
-2	-1
-2	0
-2	1



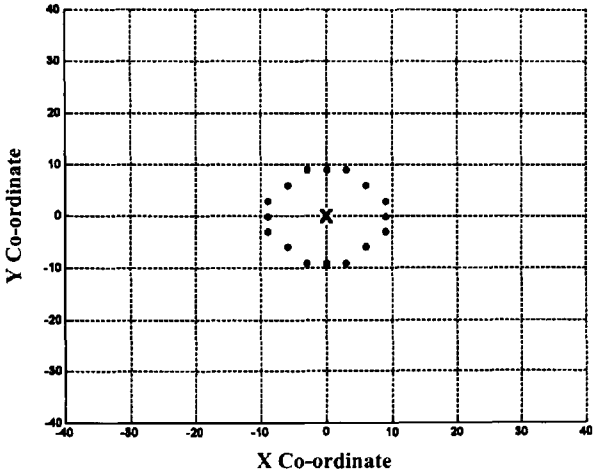
Donut

X CO-ORDINATE	Y CO-ORDINATE
-1	3
0	3
1	3
2	2
3	1
3	0
3	-1
2	-2
1	-3
0	-3
-1	-3
-2	-2
-3	-1
-3	0
-3	1
-2	2



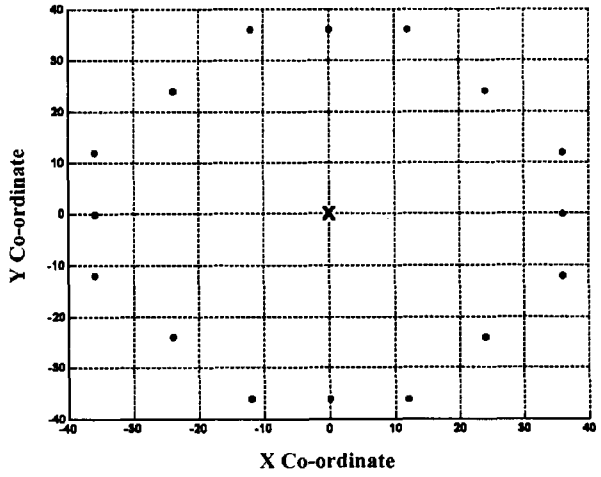
Huge nut

X CO-ORDINATE	Y CO-ORDINATE
-3	9
0	9
3	9
6	6
9	3
9	0
9	-3
6	-6
3	-9
0	-9
-3	-9
-6	-6
-9	-3
-9	0
-9	3
-6	6



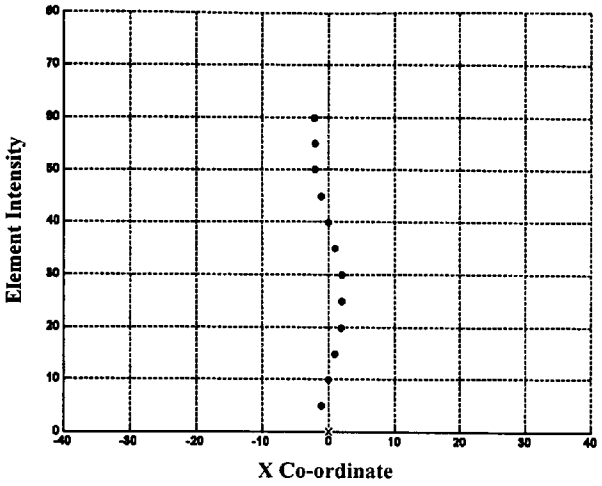
Giant nut

X CO-ORDINATE	Y CO-ORDINATE
-12	36
0	36
12	36
24	24
36	12
36	0
36	-12
24	-24
12	-36
0	-36
-12	-36
-24	-24
-36	-12
-36	0
-36	12
-24	24

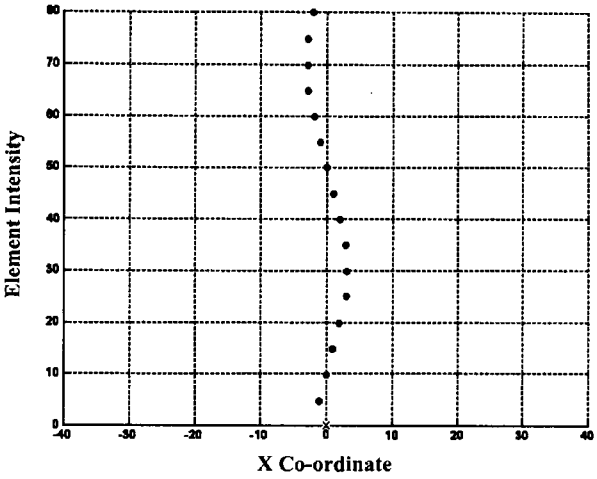


Sparse annular spiral collection (side-view) - locus (origin) represented by centre point x

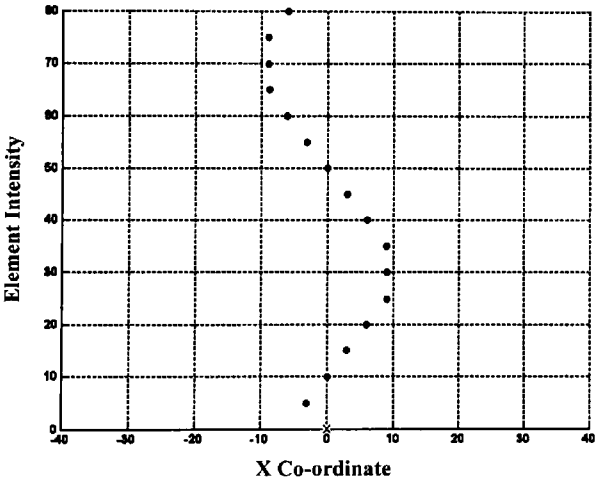
Small nut spiral



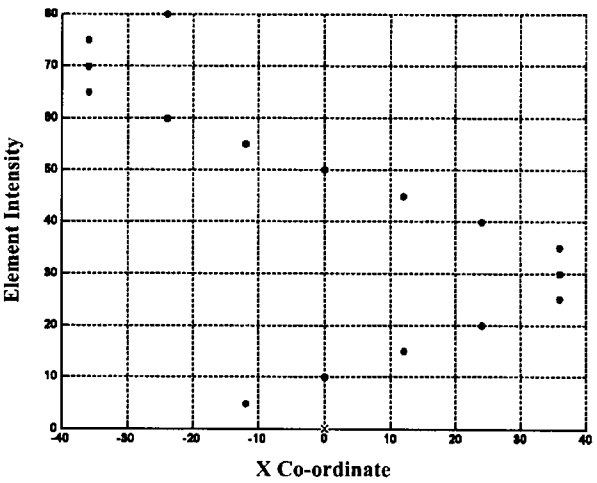
Donut spiral



Huge nut spiral

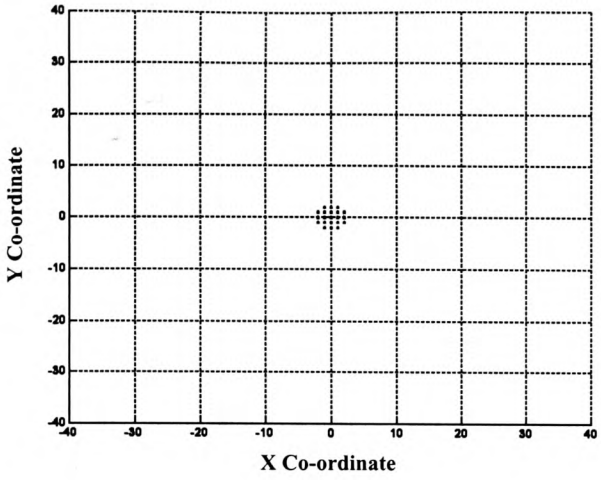


Giant nut spiral

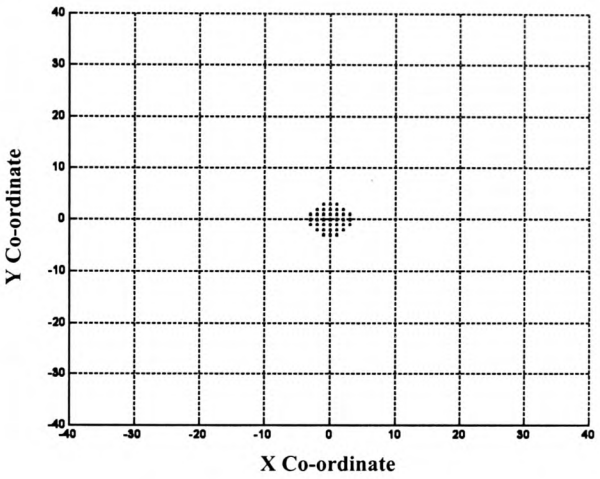


Filled template collection

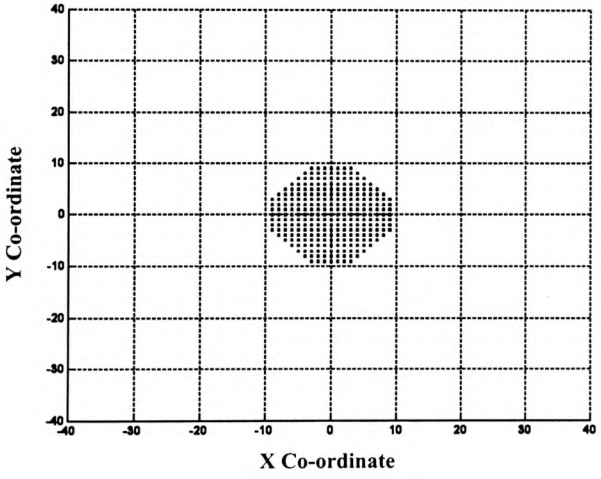
Small nut filled



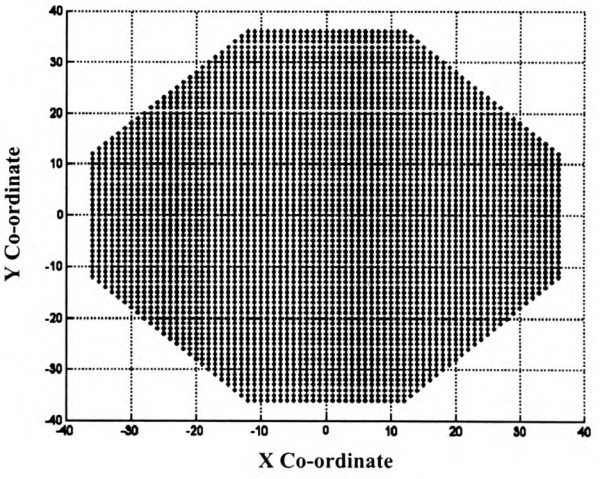
Donut filled



Huge nut Filled

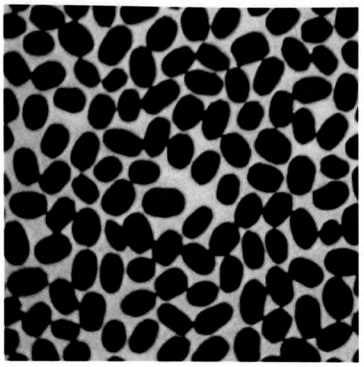


Giant nut filled

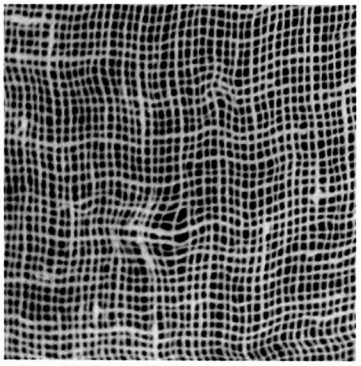


Appendix 4

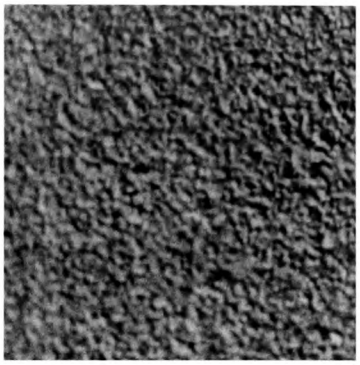
Texture samples drawn from the Brodatz set [146]



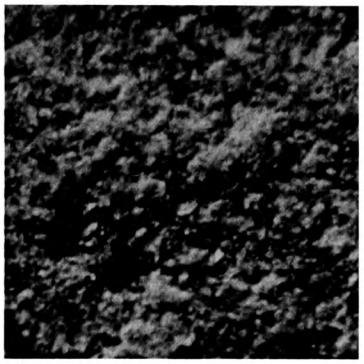
Beans



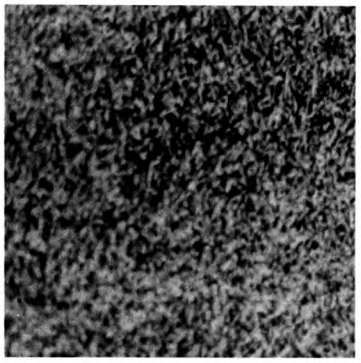
Burlap



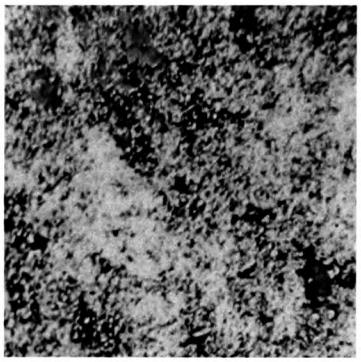
Cork



Fieldsto



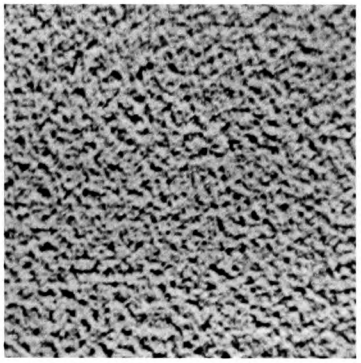
Grass



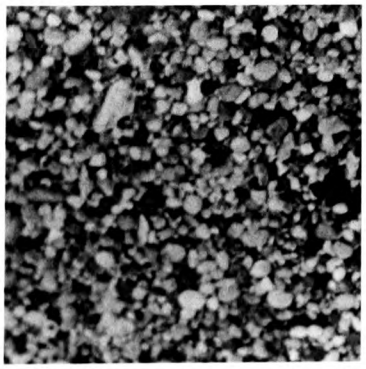
Ice

Marble

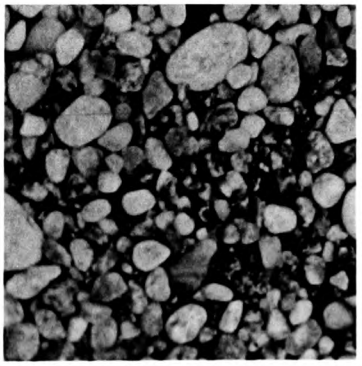
Mica



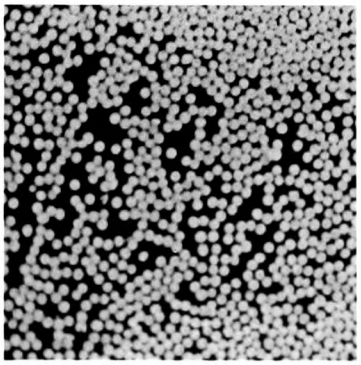
Paper



Peb54



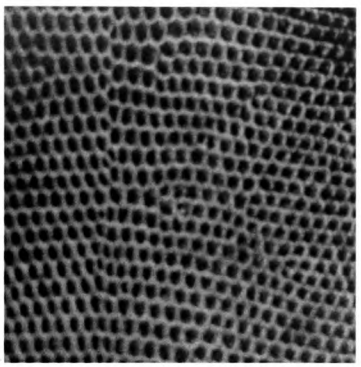
Pebbles 2



Pellets

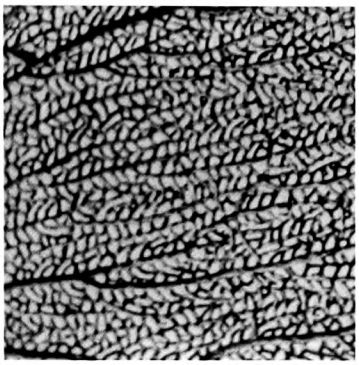
Pig skin

Raffia

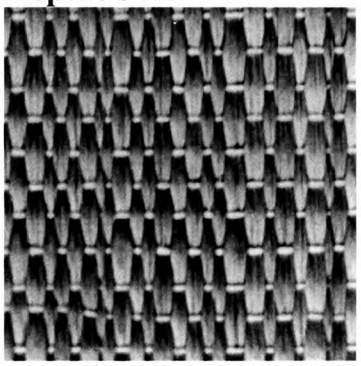


Reptile skin

Rice paper



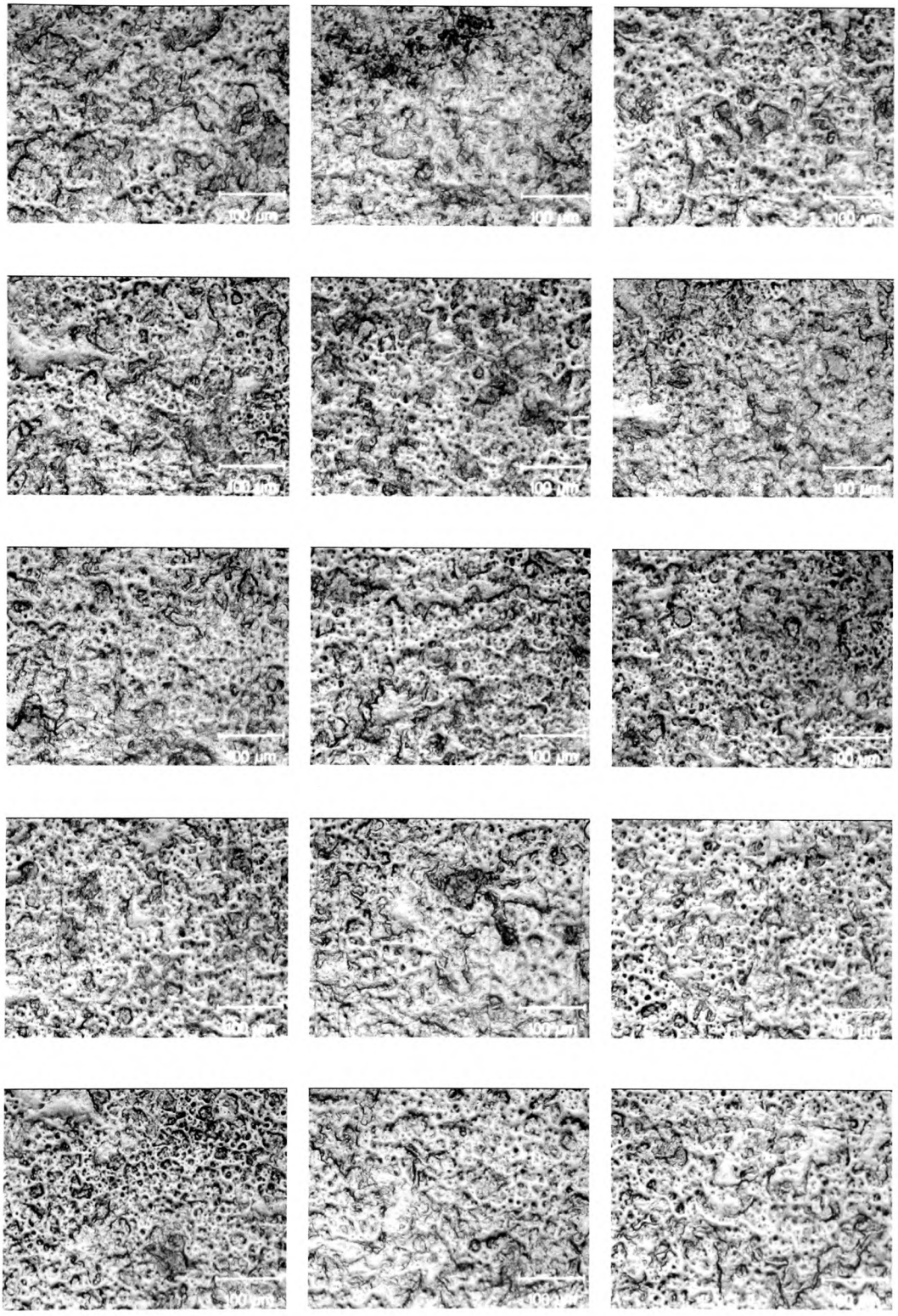
Sea fan



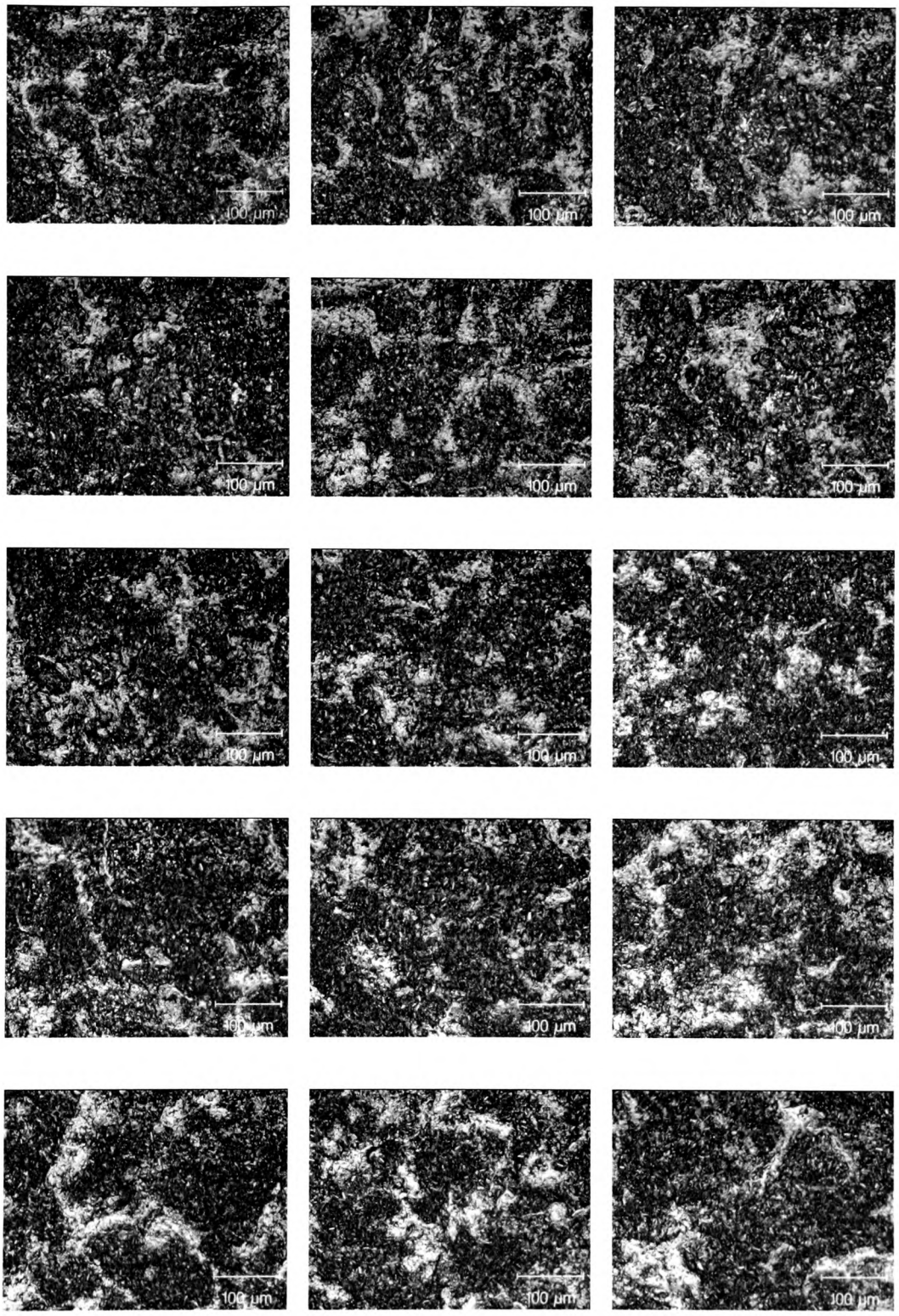
Straw

Appendix 5

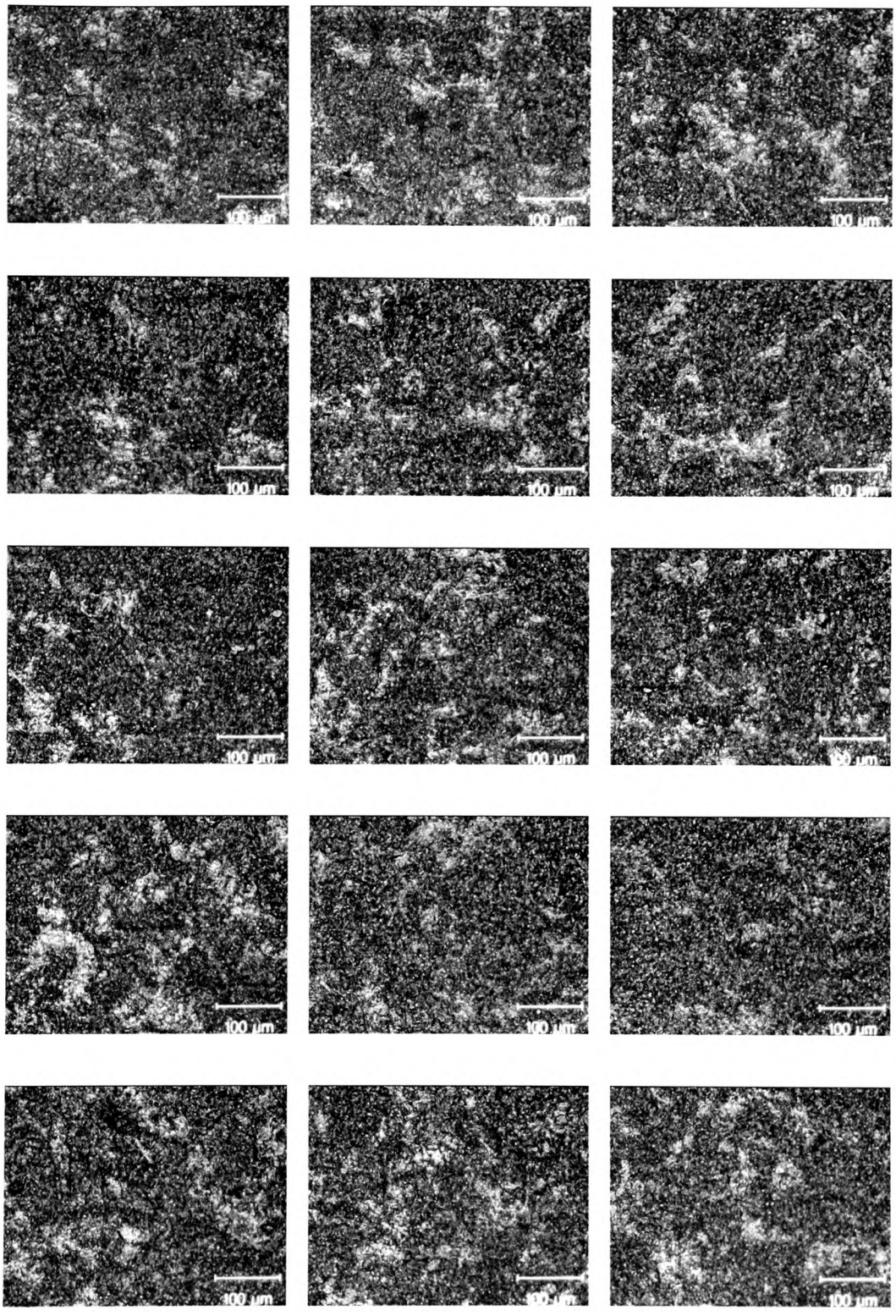
x200 image samples drawn from each annealing class



Under-alloyed



Optimum-alloyed condition



Over-alloyed

Appendix 6

Work published during the project.

"Use of Rank-Conditioned Morphological Operators for Texture Classification", Proceedings of the IEE Sixth International Conference on Image Processing and its Applications, pp. 618-622, Jul. 1997.

Use of Rank-Conditioned Morphological Operators for Texture Classification

S J Rees, S Woodham, L S Dooley

University of Glamorgan, UK

Introduction

The use of morphological decomposition and classification for texture has received considerable attention [1,2]. The basis of mathematical morphology is the study of the geometrical structure in an image through the use of simple probe forms called structuring elements. Methods have been applied to obtaining structural decomposition using a variety of sequenced applications of morphological filters. These use a sequential elimination of increasing sizes of information from the original image through application of morphological openings. This results in a multidimensional classification of the original texture based on the residual images, in [2] using grey level means for each point in the component images as texture features.

A different formulation has been implemented, using a rank-conditioned morphological implementation [3] to generate the required classification features. The features are based on the rank required for each point in the image to remain part of the rank conditionally eroded result based on the structuring element used. The structuring elements are chosen as increasing sizes of the same basic shape, in the case presented here using a sparse ring or annulus as the shape outline. Intensity variations are considered as part of the feature set through the use of increasing intensity gradients of ring.

The resulting feature vectors are used as discriminants of the texture under analysis.

Rank-Conditioned Morphological Filters

In accordance with the usual morphological filter implementations, the probe structuring element is placed at the top surface each point in the image function $f(x)$. The number of components of the structuring element not contained over the element window are calculated. This, effectively, is the minimum rank condition that would be applied to permit the point to belong to the result unaltered if a rank conditioned erosion were undertaken. This is derived from a use of rank conditioning as a means of reducing the perfection of containment required by formal set erosion. The ranked values are used to generate a characteristic signature of the texture under analysis. Note that this is not a decomposition operation, but an iterative application of different sized structuring elements to generate a multidimensional feature vector.

Consider the top surface form of greyscale erosion (see Sternberg [4], or Haralick et al [5],):

$$f \Theta k(x) = T[U[f] \Theta U[k]]$$

This may be physically interpreted as:

$$f \Theta k(x) = \min_{z \in k} \{f(x+z) - k(z)\}$$

Now, for any point $(x_i) \in f(x) | (x_i + z) \in D_f$,

if $([U[f(x_i+z)] \geq [U[f(x_i)+k(z)]) \forall z \in k(z)$

then $(f \Theta k)(x_i) = f(x_i)$

In other words, any point in the original image belongs to the eroded result provided it meets this criterion. In order to relax the criterion, we must allow the point value $f(x_i)$ to belong to the solution if components fail to meet the criterion for formal erosion, i.e. if $([U[f(x_i+z)] \geq [U[f(x_i)+k(z)])$ for some proportion of $z \in k(z)$ rather than for its whole. If the test is to be general, rather than directional or for a specific fragment of the probe, then the metric for assessing coverage should simply be the count of how many components of the probe set are contained.

In effect our measure becomes:

$(f @ k)(x) = R(x), J$, where

$R(x) = \{r(x): N, x \in f(x), (x+z \in D_f)|$

$r = \sum_{z \in k(z)} m$, $m=1$ where $(f(x+z) \geq (k(z) + f(x))$;

$m = 0$ otherwise}

A $J(x)$ set may be generated for each locus in $f(x)$ (and hence $R(x)$), and will contain a number of elements up to the number of elements

in $k(z)$, indicating the structural information missing from $f(x)$ to permit complete containment of $k(z)$. It is primarily of use in recognition model evaluation from set occupancy, and is not used in the texture evaluation discussed here.

$$J(x) = \{J_r\} \forall r \in R(x)$$

$$= \{j : N \mid j = k(z) \text{ where } f(x+z) = 0,$$
$$j = k(z) - f(r+z) \text{ where } f(r+z) < k(z),$$
$$\forall z \in k(z)\} \forall r \in R(x)$$

The analysis is then formed as a histogram, H, where, over a sample window size MxN$\in f(x,y)$, the individual histogram values are:

$$H_i = \sum_{M,N} r(x,y) = i, i <= CARD(k(z))$$

The set of characteristic histograms, **H**, is the feature classifier.

$$H_N = \{\{H\}_n\}, n = \text{no of probe structuring elements applied}$$

Note that although the sample window applied is square, the resulting shape of area analysed depends, additionally, on the shape of the probe structuring elements and does overlap the perimeter of the window area in its analysis.

Experimental Results

The structuring elements used were a set of increasing sizes of a sparse annular templates. These were applied in turn to the acquired texture images and the resulting feature histograms built, and applied to the classification of other samples of textural data. The textures used were drawn from the natural texture examples in the Brodatz [6] set of samples, and also from sample images of galvannealed steel with different degrees of annealing. The Brodatz set textures included D3, D4, D5, D9, D15, D54, D57, D62, D67, D92, D98, D103 Typical examples of natural textures are shown below in figures (1) and (2):

Figure 1: Texture Sample Beans

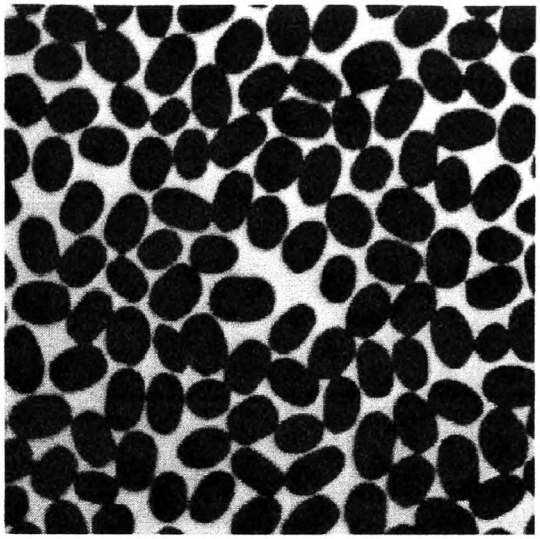
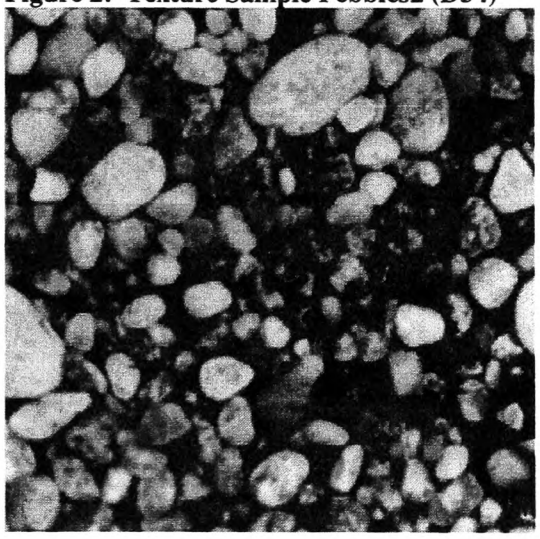


Figure 2: Texture Sample Pebbles2 (D54)



The effects of sample window size on the histograms obtained was evaluated. The resulting classifiers applied to nineteen texture samples drawn from the Brodatz set and the separations of the feature vectors calculated. The images were based on a 256x256, 8 bit deep sampling. The results shown in Table (1) below were obtained using a sparse annular template. Classification was successfully obtained for adequate samples of texture with five sizes and three intensity gradients of template, yielding a fifteen dimensional evaluation vector for each sample. In this example, the fifteen evaluation feature vectors were found to be redundant for the samples taken. It should be noted that, in

this case, the samples were taken as sub-images of each texture. Complete isolation of the texture within the system world model is claimed where the in-sample scatter is less than the between sub-samples worst case scatter (the data was then experimentally verified).

Table (1): Effect of Window Size and Number of Vectors on Texture Discrimination

Window Size (M=N)	Number of Feature Vectors	Unique Discrimination of All Texture Samples %	Average Level of Discrimination %	Lowest Level of Discrimination of Textures %
48	15	100	100	100
96	15	100	100	100
48	3	26	80	50
96	3	100	100	100

The feature extraction method was then applied to a natural, random texture sample, using 380 samples of galvannealed steel in five known, calibrated coating conditions. These images were acquired at a x500 magnification, and were subject in some cases to depth of field problems causing partial blurring of the acquired image as is shown in figure 3. As can be seen from the results of table 2, there is a strong correlation between the coating condition and the optical texture as it was measured, but the classification scheme is by no means perfect in its present form.

Table 2: Classification of Galvanneal Samples

Window Size (MxN)	No of Vectors Used	% Assigned to Correct Class (of 380 Samples, 5 Classes)
400x400	15	75
200x200	15	65
400x400	3	60

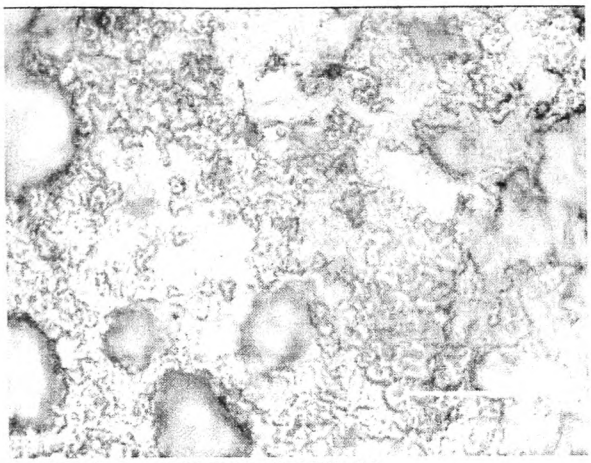
Noise Performance

Salt and pepper noise was added to the Brodatz set texture samples, and the algorithms run to classify them. In the context of the rank of the containment of the probe sets, this impulsive noise produces limited distortion of the histogram, but causes the raising of the outliers in the data (corresponding to no containment and full containment at a point). The algorithms were adjusted to ignore these outliers, producing a significant improvement in the performance of the classification scheme.

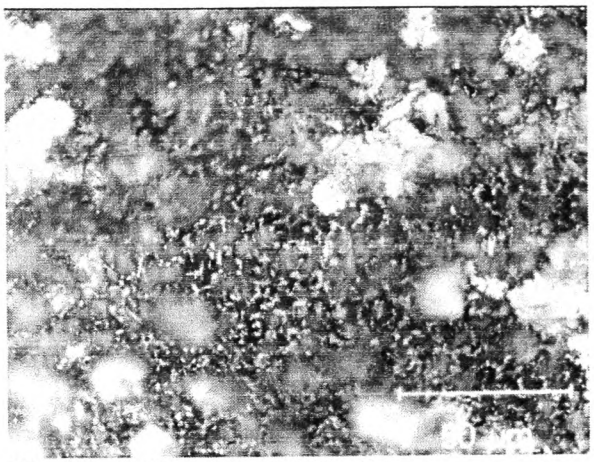
Table 3: Noise Performance of Classification Based on the Use of 15 Feature Vectors

% Noise	Window Size	% Correctly Classified (of 19 Texture Samples)	% Correctly Classified on Removal of Data Outliers
0	150x150	100	100
1	150x150	90	95
2	150x150	79	85
3	150x150	63	70
4	150x150	52	60

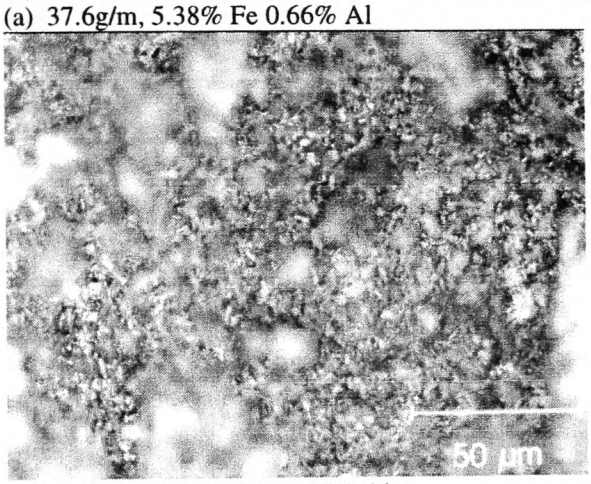
Figure 3: Galvanneal Samples, 512x512 Images at 8 Bit Resolution



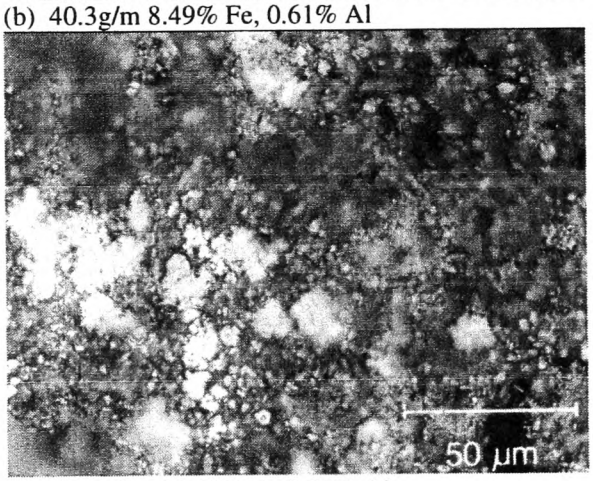
(a) 37.6g/m, 5.38% Fe 0.66% Al



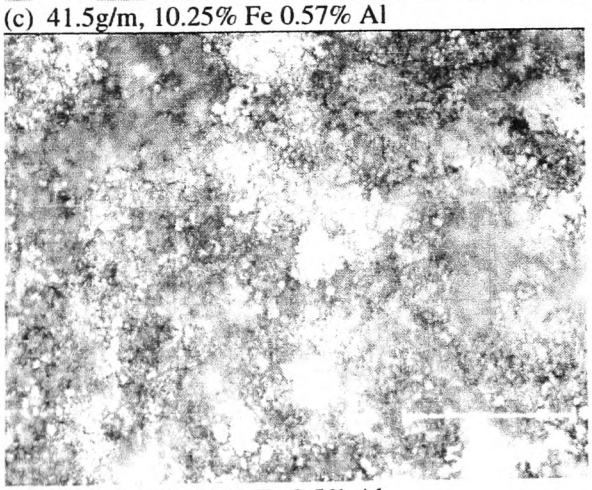
(b) 40.3g/m 8.49% Fe, 0.61% Al



(c) 41.5g/m, 10.25% Fe 0.57% Al



(d) 42.7g/m, 11.48% Fe 0.57% Al



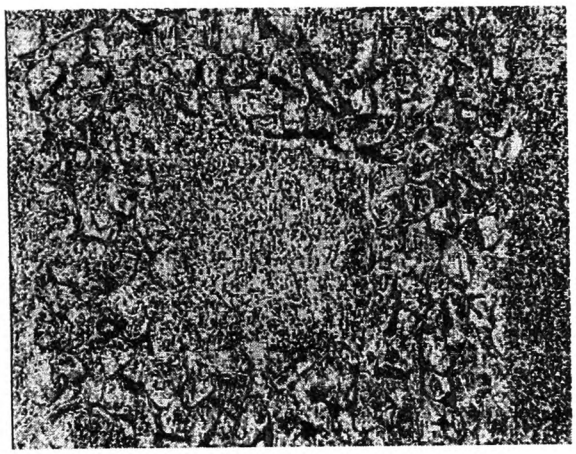
(e) 51.2g/m, 13.64 % Fe 0.5% Al

Boundary Detection

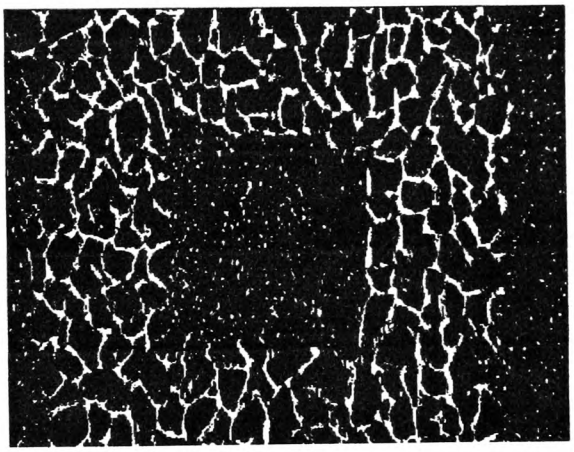
In order to locate boundaries between textures, it is necessary to assign pixels in close proximity to the boundary to particular texture regions. The larger the probe structuring element is about the point, the more likely it is to overlap the "other" texture region. The approach used was to adopt a 96x96 pixel window size, and classify the regions initially on this basis. Where the classification metric exceeds the permitted variation, typically near a texture boundary, then the larger sizes of template were eliminated and the window size halved, down to a limit of 12

pixels window size. All remaining point are assigned on a nearest neighbour basis. With test data of the form shown in figure 4 below, the pixel classification was found to be accurate down to the minimum window size, and the system classified between 30% and 70% of the remaining pixels correctly, dependant upon the chosen boundary shape.

Figure (4): Textured Object and Background Enhancement



(i) Gravel on Chipboard



(ii) Textures Enhanced Separation

Conclusion

The texture analysis method developed has been successfully applied to a variety of natural textures. It is based on extraction of integer data, and the histogram feature data extraction is capable of implementation in hardware, based on a shift and compare algorithm on a pixel by pixel basis, comparison being undertaken over the chosen template shape size. Subject to the replication of the shape size area, this is highly amenable to parallel implementation down to the single pixel level.

Comparison with the work of Wang and Haese-Coat[2] in terms of noise performance and discrimination indicate a less effective result. This is due in part to the non-optimal classification method adopted, and is compensated by lower computational intensity.

References:

1. Toet A., 1989, "A morphological pyramidal image decomposition", Pattern Recognition Letters, No 9, pp255-261.

2. Wang D, Haese-Coat V, Bruno A, and Ronsin J., 1993, "Texture Classification and Segmentation Based on Iterative Morphological Decomposition", Journal of Visual Communication and Image Representation, No 4, pp197-214.

3. Rees S J, Jones B F, "Conditional Morphological Operators for Direct Feature Extraction and Enhancement", 5th International Conference on Image Processing and its Applications, IEE Conference Publication No. 410, Edinburgh, July 1995, pp747-751.

4. Sternberg S R, "Greyscale Morphology", Computer Vision, Graphics and Image Processing, Vol. 35, pp333-355, 1986.

5. Haralick R, Sternberg S R, Zhuang X, "Image Analysis using Mathematical Morphology", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, No. 4, July 1987.

6. Brodatz P, Textures, Dover, New York.

7. Wei Li, Haese-Coat V, "Composite Morphological Filters in Multiresolution Morphological Decomposition", 5th International Conference on Image Processing and its Applications, IEE Conference Publication No. 410, Edinburgh, July 1995, pp752-756.